



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 145028

TO: Bao-Qun Li  
Location: rem/3d24/3c18  
Art Unit: 1648  
Wednesday, February 23, 2005  
  
Case Serial Number: 09/664363

From: Deirdre Arnold  
Location: Biotech-Chem Library  
REM 1A64  
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

### Search Notes

*Please feel free to contact me if you have any questions or would like to amend the search.*

Thank you for using STIC services.

Regards,  
Deirdre Arnold

**This Page Blank (uspto)**

2

145028

my

STIC-Biotech/ChemLib

From: Li, Bao-Qun  
Sent: Monday, February 14, 2005 8:42 AM  
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF SEQ ID NO; 19 IN APPLICATION SN.  
09,664,363. THANKS.

Bao Qun Li M.D  
TC 1600  
Art Unit 1648  
Tel. 517-272-0904  
REM, 3C18  
Rm. 3D24

RECEIVED  
FEB 14 2005  
STIC

\*\*\*\*\*  
STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*  
Type of Search

NA Sequence: # \_\_\_\_\_  
AA Sequence :# \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*  
Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**





# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 145027**

**TO: Bao-Qun Li**  
**Location: rem/3d24/3c18**  
**Art Unit: 1648**  
**Sunday, February 20, 2005**

**Case Serial Number: 09/664363**

**From: Toby Port**  
**Location: Biotech-Chem Library**  
**REM-1A59**  
**Phone: 571-272-2523**

**toby.port@uspto.gov**

### **Search Notes**

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Toby Port

**This Page Blank (uspto)**



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



**This Page Blank (uspto)**

①

STIC-Biotech/ChemLib

145027

1/1/05

From: Li, Bao-Qun  
Sent: Monday, February 14, 2005 8:41 AM  
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF SEQ ID NO; 5 IN APPLICATION SN.  
09,664,363. THANKS.

Bao Qun Li M.D  
TC 1600  
Art Unit 1648  
Tel. 517-272-0904  
REM, 3C18  
Rm. 3D24

RECEIVED  
FEB 14 2005  
STIC

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA Sequence: # \_\_\_\_\_  
AA Sequence: # \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 20, 2005, 22:06:50 / Search time 4114 Seconds  
(without alignments)  
9822.958 Million cell updates/sec

Title: US-09-664-363-5  
Perfect score: 834  
Sequence: 1 AGAATAACCAACGTAACAC.....TCTGGCATCTGTTCCCG 834

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hc:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	834	100.0	834	6	A28135 PT-NANBH mr
2	834	100.0	834	6	A32186 NANBH PT P
3	834	100.0	834	6	AR144034 Sequence 1
4	804.8	96.5	2790	6	A76569 Sequence 1
5	804.8	96.5	3372	6	A76594 Sequence 26
6	802.6	96.2	2116	6	A28155 PT-NANBH mr
7	802.6	96.2	2116	6	A32202 NANBH PT P
8	802.6	96.2	2116	6	AR144050 Sequence 3
9	801.6	96.1	831	6	A76571 Sequence 3
10	748.2	89.7	933	14	HCV132997 Hepatitis C
11	746.6	89.5	1595	14	HCPCNS1SPF Hepatitis C
12	746.6	89.5	9379	14	AF207768 Hepatitis C
13	745	89.3	9379	14	AF165051 Hepatitis C
14	745	89.3	9379	14	AF165052 Hepatitis C
15	745	89.3	9410	14	HPCK1R2 Hepatitis C
16	743.4	89.1	3286	14	AB008441 Hepatitis C
17	743.4	89.1	3286	14	AB008443 Hepatitis C
18	743.4	89.1	9386	14	AF165055 Hepatitis C
19	743.4	89.1	9386	14	AF165056 Hepatitis C

20	743.4	89.1	9418	14	HCV132996 Hepatitis C
21	741.8	88.9	1635	14	M74809 Hepatitis C
22	741.8	88.9	2672	14	HCPCNS1SPF Hepatitis C
23	741.8	88.9	3296	14	AB008442 Hepatitis C
24	741.8	88.9	9373	14	AF208024 Hepatitis C
25	741.8	88.9	9374	14	AF207753 Hepatitis C
26	741.8	88.9	9573	14	AB049090 Hepatitis C
27	741.8	88.9	9585	14	AB049095 Hepatitis C
28	740.2	88.8	9379	14	AF207761 Hepatitis C
29	740.2	88.8	9379	14	AF207766 Hepatitis C
30	740.2	88.8	9379	14	AF207773 Hepatitis C
31	740.2	88.8	9410	14	HCPCNS2 Hepatitis C
32	738.6	88.6	9402	6	E05027 Hepatitis C
33	738.6	88.6	9431	6	HCU45476 Hepatitis C
34	737	88.4	2433	6	A48711 Sequence 49
35	737	88.4	2433	6	AR157351 Sequence
36	737	88.4	2433	6	AX452798 Sequence
37	737	88.4	2433	6	AX685050 Sequence
38	737	88.4	2433	6	AX794892 Sequence
39	737	88.4	2672	14	HCVUK5 Hepatitis C
40	737	88.4	9033	14	DB9872 Hepatitis C
41	737	88.4	9369	14	AF165053 Hepatitis C
42	737	88.4	9426	14	AB049087 Hepatitis C
43	737	88.4	9547	14	AB049091 Hepatitis C
44	737	88.4	9587	14	AF333324 Hepatitis C
45	735.4	88.2	1195	14	HCPCOPR L38318 Hepatitis C

## ALIGNMENTS

RESULT 1  
A28135  
LOCUS A28135 834 bp DNA linear PAT 07-JUN-1995  
DEFINITION PT-NANBH mRNA fragment from patent GB2239245.  
ACCESSION A28135  
VERSION A28135.1 GI:1248620  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
other sequences; artificial sequences.  
REFERENCE  
1 (bases 1 to 834)  
AUTHORS  
JOURNAL  
FEATURES  
source  
Post-translational non-A non-B hepatitis viral polypeptides  
Location/Qualifiers  
1..834  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/clone="BR11"  
<1..>834  
/codon\_start=1  
/transl\_table=11  
/product="PT-NANBH polypeptide"  
/protein\_id="CAA01927.1"

## CDS

ORIGIN  
Query Match 100.0%; Score 834; DB 6; Length 834;  
Best Local Similarity 100.0%; Pred. No. 5.2e-193;  
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 AGAATAACCAACGTAACACCTCCGCCACAGAGCTCAGGTTCCCGCGGTGGT 60  
|||||  
1 AGAATAACCAACGTAACACCTCCGCCACAGAGCTCAGGTTCCCGCGGTGGT 60  
|||||

DB 1 AGAATAACCAACGTAACACCTCCGCCACAGAGCTCAGGTTCCCGCGGTGGT 60  
|||||

```

QY 61 CAGATCGTTGGTGAAGTTTAACTGTTGCGCGCGCAAGGCGCCCAAGTTGGGTGCGCGG 120
Db 61 CAGATCGTTGGTGAAGTTTAACTGTTGCGCGCGCAAGGCGCCCAAGTTGGGTGCGCGG 120
QY 121 ACTAGGAAGACTTCCGAGCGGTGCGCAACTCGTGGAAAGGCGCAACAATTATCCCAAGGT 180
Db 121 ACTAGGAAGACTTCCGAGCGGTGCGCAACTCGTGGAAAGGCGCAACAATTATCCCAAGGT 180
QY 181 CGCGAGCCCGAGGCGAAGGCGCTGAGCGTCAAGCCCGGGTACCTTTGGCCCTCATGGAAC 240
Db 181 CGCGAGCCCGAGGCGAAGGCGCTGAGCGTCAAGCCCGGGTACCTTTGGCCCTCATGGAAC 240
QY 241 GAGGCGATGGGTGGGCGAGAGTGGCTCTGTCAACCCCGGTGCTCCCGGCTTGTGGGCG 300
Db 241 GAGGCGATGGGTGGGCGAGAGTGGCTCTGTCAACCCCGGTGCTCCCGGCTTGTGGGCG 300
QY 301 CCACCTGACCCCGGGGTGAGTGGGTAAATTGGGTAAGTATGATGATACCTTCAATGC 360
Db 301 CCACCTGACCCCGGGGTGAGTGGGTAAATTGGGTAAGTATGATGATACCTTCAATGC 360
QY 361 GGGCTTCGCGCACTCATGGGGGTAACTTCGCTCGTGGCGGCTCCCTTAAGGGCGCTGCG 420
Db 361 GGGCTTCGCGCACTCATGGGGGTAACTTCGCTCGTGGCGGCTCCCTTAAGGGCGCTGCG 420
QY 421 AGGCGCTTGGCGCATAGCGGTCCGGGTTCTGGAGAGACGGGTGATATATGCAAGGAAT 480
Db 421 AGGCGCTTGGCGCATAGCGGTCCGGGTTCTGGAGAGACGGGTGATATATGCAAGGAAT 480
QY 481 TTACCCGGTGGCTCTTCTCTATCTTCCCTTGGGCTTGGCTGCTGCTTGGCAATTTCCA 540
Db 481 TTACCCGGTGGCTCTTCTCTATCTTCCCTTGGGCTTGGCTGCTGCTTGGCAATTTCCA 540
QY 541 GCTTCGCGTTATGAAGTGGCGAAGTGTCCGGGATCTACATGTACAGAACGATTGCTCC 600
Db 541 GCTTCGCGTTATGAAGTGGCGAAGTGTCCGGGATCTACATGTACAGAACGATTGCTCC 600
QY 601 AACTAAGCATGTGTAGAGAGACAGCGACATGATATGACACACCCCGGGGTGTGCTCC 660
Db 601 AACTAAGCATGTGTAGAGAGACAGCGACATGATATGACACACCCCGGGGTGTGCTCC 660
QY 661 TGTGTCGGGAGAGGTAATTCCTCCGCTGCTGGGTAAGGCTCACTCCACGCTCGGCGG 720
Db 661 TGTGTCGGGAGAGGTAATTCCTCCGCTGCTGGGTAAGGCTCACTCCACGCTCGGCGG 720
QY 721 AAGGACGCGCAGCATCCCACTGCGACAAATACAGACGCGCATGATTGCTGTTGGGCG 780
Db 721 AAGGACGCGCAGCATCCCACTGCGACAAATACAGACGCGCATGATTGCTGTTGGGCG 780
QY 781 GCTGCTTCTGCTCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCCCG 834
Db 781 GCTGCTTCTGCTCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCCCG 834

```

```

RESULT 2
A32186 834 bp RNA linear PART 10-DEC-1996
LOCUS NAMHV PT polypeptide fragment.
DEFINITION A32186
ACCESSION A32186
VERSION A32186.1 GI:1926542
KEYWORDS
SOURCE Non-A, non-B hepatitis virus
ORGANISM Non-A, non-B hepatitis virus
VIRUSES; unclassified viruses.
REFERENCE 1 (bases 1 to 834)
AUTHORS Patent: FR 2655990-A 5 21-JUN-1991;
JOURNAL Location/Qualifiers
FEATURES
source
1..834
/organism="Non-A, non-B hepatitis virus"
/mol_type="unassigned RNA"
/db_xref="taxon:12440"
/clone="BR11"
/clone_lib="lambda gt11"

```

```

CDS
<1..>834
/codon_start=1
/product="partial NAMHV-PT polypeptide"
/protein_id="CAA02090.1"
/db_xref="GI:1926543"
/translation="RTKRTNTNLRPODVPFPGGQIVGGVLLPBRGRRLRATKRT
SERQPRRQPIKARQPEBRARARQPEPPLVGNBEMGAMHLSRGRPSMPT
DPRRSRLNGLVLDLTCGFADSHGVSHARRSLRGARALAHGVRLVLDGNATGN
LPQGSFSLFLALISCLTI PASAEVRNVSQIYHTVNDSCSIYETADMIHTPGC
VPCVREGNSRCWVALPTPLAAKDASIPATIRRHVDLLVGAFAFSSAMVYGDLCGSV
FP"
ORIGIN
Query Match 100.0%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 5,2e-193;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AGAAAAACCAACGTAACCAACCACTCCGCCACAGACGTCAAGGTTCCCGGGCGGTGT 60
Db 1 AGAAAAACCAACGTAACCAACCACTCCGCCACAGACGTCAAGGTTCCCGGGCGGTGT 60
QY 61 CAGATCGTTGGTGAAGTTTAACTGTTGCGCGCGCAAGGCGCCCAAGTTGGGTGCGCGG 120
Db 61 CAGATCGTTGGTGAAGTTTAACTGTTGCGCGCGCAAGGCGCCCAAGTTGGGTGCGCGG 120
QY 121 ACTAGGAAGACTTCCGAGCGGTGCGCAACTCGTGGAAAGGCGCAACAATTATCCCAAGGT 180
Db 121 ACTAGGAAGACTTCCGAGCGGTGCGCAACTCGTGGAAAGGCGCAACAATTATCCCAAGGT 180
QY 181 CGCGAGCCCGAGGCGAAGGCGCTGAGCGTCAAGCCCGGGTACCTTTGGCCCTCATGGAAC 240
Db 181 CGCGAGCCCGAGGCGAAGGCGCTGAGCGTCAAGCCCGGGTACCTTTGGCCCTCATGGAAC 240
QY 241 GAGGCGATGGGTGGGCGAGAGTGGCTCTGTCAACCCCGGTGCTCCCGGCTTGTGGGCG 300
Db 241 GAGGCGATGGGTGGGCGAGAGTGGCTCTGTCAACCCCGGTGCTCCCGGCTTGTGGGCG 300
QY 301 CCACCTGACCCCGGGGTGAGTGGGTAAATTGGGTAAGTATGATGATACCTTCAATGC 360
Db 301 CCACCTGACCCCGGGGTGAGTGGGTAAATTGGGTAAGTATGATGATACCTTCAATGC 360
QY 361 GGGCTTCGCGCACTCATGGGGGTAACTTCGCTCGTGGCGGCTCCCTTAAGGGCGCTGCG 420
Db 361 GGGCTTCGCGCACTCATGGGGGTAACTTCGCTCGTGGCGGCTCCCTTAAGGGCGCTGCG 420
QY 421 AGGCGCTTGGCGCATAGCGGTCCGGGTTCTGGAGAGACGGGTGATATATGCAAGGAAT 480
Db 421 AGGCGCTTGGCGCATAGCGGTCCGGGTTCTGGAGAGACGGGTGATATATGCAAGGAAT 480
QY 481 TTACCCGGTGGCTCTTCTCTATCTTCCCTTGGGCTTGGCTGCTGCTTGGCAATTTCCA 540
Db 481 TTACCCGGTGGCTCTTCTCTATCTTCCCTTGGGCTTGGCTGCTGCTTGGCAATTTCCA 540
QY 541 GCTTCGCGTTATGAAGTGGCGAAGTGTCCGGGATCTACATGTACAGAACGATTGCTCC 600
Db 541 GCTTCGCGTTATGAAGTGGCGAAGTGTCCGGGATCTACATGTACAGAACGATTGCTCC 600
QY 601 AACTAAGCATGTGTAGAGAGACAGCGACATGATATGACACACCCCGGGGTGTGCTCC 660
Db 601 AACTAAGCATGTGTAGAGAGACAGCGACATGATATGACACACCCCGGGGTGTGCTCC 660
QY 661 TGTGTCGGGAGAGGTAATTCCTCCGCTGCTGGGTAAGGCTCACTCCACGCTCGGCGG 720
Db 661 TGTGTCGGGAGAGGTAATTCCTCCGCTGCTGGGTAAGGCTCACTCCACGCTCGGCGG 720
QY 721 AAGGACGCGCAGCATCCCACTGCGACAAATACAGACGCGCATGATTGCTGTTGGGCG 780
Db 721 AAGGACGCGCAGCATCCCACTGCGACAAATACAGACGCGCATGATTGCTGTTGGGCG 780
QY 781 GCTGCTTCTGCTCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCCCG 834
Db 781 GCTGCTTCTGCTCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCCCG 834

```



```

RESULT 3
LOCUS      AR144034                      834 bp      DNA      linear      PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6210675.
ACCESSION  AR144034
VERSION     AR144034.1  GI:15105901
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   Unidentified.
AUTHORS     1 (bases 1 to 834)
            Higfield, P., Edmund., Rodgers, B., Colin., Tedder, R., Seton. and
            Barbara, J. Anthony, James.
TITLE       PT-NANB hepatitis polypeptides
JOURNAL     Patent: US 6210675-A 5 03-APR-2001;
            Location/Qualifiers
FEATURES
            1..834
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 5,2e-193;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAAAAACCAACGTAACCAACCACTCCGCCACAGAGCTCAGGTTCCCGGGCGGTGT 60
DB      1 AGAAAAACCAACGTAACCAACCACTCCGCCACAGAGCTCAGGTTCCCGGGCGGTGT 60
QY      61 CAGATCGTTGGTGAAGTTTACCTGTTGCGCGGAGGAGGCGCCAGTTGGTGTGCGGCG 120
DB      61 CAGATCGTTGGTGAAGTTTACCTGTTGCGCGGAGGAGGCGCCAGTTGGTGTGCGGCG 120
QY      121 ACTAGAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGAACCTATCCCAAGGCT 180
DB      121 ACTAGAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGAACCTATCCCAAGGCT 180
QY      121 ACTAGAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGAACCTATCCCAAGGCT 180
DB      121 ACTAGAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGAACCTATCCCAAGGCT 180
QY      181 CGCGACCCGAGAGGCGAGGCTTGCTGAGCCCGGGTACCCTTGGCCCTCTATGGCAC 240
DB      181 CGCGACCCGAGAGGCGAGGCTTGCTGAGCCCGGGTACCCTTGGCCCTCTATGGCAC 240
QY      181 CGCGACCCGAGAGGCGAGGCTTGCTGAGCCCGGGTACCCTTGGCCCTCTATGGCAC 240
DB      181 CGCGACCCGAGAGGCGAGGCTTGCTGAGCCCGGGTACCCTTGGCCCTCTATGGCAC 240
QY      241 GAGGGGATGGGGGCGAGGATGGCTCCCTGTCACCCCGGGCTCCGGCCTATTGGGGG 300
DB      241 GAGGGGATGGGGGCGAGGATGGCTCCCTGTCACCCCGGGCTCCGGCCTATTGGGGG 300
QY      241 GAGGGGATGGGGGCGAGGATGGCTCCCTGTCACCCCGGGCTCCGGCCTATTGGGGG 300
DB      241 GAGGGGATGGGGGCGAGGATGGCTCCCTGTCACCCCGGGCTCCGGCCTATTGGGGG 300
QY      301 CCCACTGACCCCGGGGAGTGCCTTAATTTGGGTAAGTCAATGATACCTCAGATGC 360
DB      301 CCCACTGACCCCGGGGAGTGCCTTAATTTGGGTAAGTCAATGATACCTCAGATGC 360
QY      301 CCCACTGACCCCGGGGAGTGCCTTAATTTGGGTAAGTCAATGATACCTCAGATGC 360
DB      301 CCCACTGACCCCGGGGAGTGCCTTAATTTGGGTAAGTCAATGATACCTCAGATGC 360
QY      361 GGGTTGCGGACTCTCATAGGGGTACATTCGCTGTCGCGCTCCCTTAGGGGCGCTGCC 420
DB      361 GGGTTGCGGACTCTCATAGGGGTACATTCGCTGTCGCGCTCCCTTAGGGGCGCTGCC 420
QY      361 GGGTTGCGGACTCTCATAGGGGTACATTCGCTGTCGCGCTCCCTTAGGGGCGCTGCC 420
DB      361 GGGTTGCGGACTCTCATAGGGGTACATTCGCTGTCGCGCTCCCTTAGGGGCGCTGCC 420
QY      421 AGGGCCCTGGCGGATGGCGCTCCGGGTTTGGAGAGAGCGGCTGAATGCAAGAGGAAT 480
DB      421 AGGGCCCTGGCGGATGGCGCTCCGGGTTTGGAGAGAGCGGCTGAATGCAAGAGGAAT 480
QY      421 AGGGCCCTGGCGGATGGCGCTCCGGGTTTGGAGAGAGCGGCTGAATGCAAGAGGAAT 480
DB      421 AGGGCCCTGGCGGATGGCGCTCCGGGTTTGGAGAGAGCGGCTGAATGCAAGAGGAAT 480
QY      481 TTACCGGTTGCTCTTCTCTATCTTCTCTGCTTGGCTTGTGTCGTTGACATTTCCA 540
DB      481 TTACCGGTTGCTCTTCTCTATCTTCTCTGCTTGGCTTGTGTCGTTGACATTTCCA 540
QY      481 TTACCGGTTGCTCTTCTCTATCTTCTCTGCTTGGCTTGTGTCGTTGACATTTCCA 540
DB      481 TTACCGGTTGCTCTTCTCTATCTTCTCTGCTTGGCTTGTGTCGTTGACATTTCCA 540
QY      541 GCTTCGCTTATAGAGTGGCAACGTGTCCGGGATCTACCATGTCAAGACATTTGCTCC 600
DB      541 GCTTCGCTTATAGAGTGGCAACGTGTCCGGGATCTACCATGTCAAGACATTTGCTCC 600
QY      541 GCTTCGCTTATAGAGTGGCAACGTGTCCGGGATCTACCATGTCAAGACATTTGCTCC 600
DB      541 GCTTCGCTTATAGAGTGGCAACGTGTCCGGGATCTACCATGTCAAGACATTTGCTCC 600
QY      601 AACTCAAGACTCTGTACAGAGAGAGAGAGATATATAGACACCCCGGGGTGTGTCGCC 660
DB      601 AACTCAAGACTCTGTACAGAGAGAGAGAGATATATAGACACCCCGGGGTGTGTCGCC 660
QY      601 AACTCAAGACTCTGTACAGAGAGAGAGAGATATATAGACACCCCGGGGTGTGTCGCC 660
DB      601 AACTCAAGACTCTGTACAGAGAGAGAGAGATATATAGACACCCCGGGGTGTGTCGCC 660
QY      661 TGTGTCCGGAGGGTAATTCCTCCGCTGCTGGGTAGCGCTCACTCCCAAGCTCGCGGCC 720
DB      661 TGTGTCCGGAGGGTAATTCCTCCGCTGCTGGGTAGCGCTCACTCCCAAGCTCGCGGCC 720

```

```

QY      721 AAGAGCGCCAGCATCCCGCACTGCGACAATACAGACGCAAGCTGATTTGCTGTGGGCG 780
DB      721 AAGAGCGCCAGCATCCCGCACTGCGACAATACAGACGCAAGCTGATTTGCTGTGGGCG 780
QY      781 GCTGCTCTCTGCTCGCTATAGCTGAGGAGATCTCTGCGAGTCTGTTTCCCG 834
DB      781 GCTGCTCTCTGCTCGCTATAGCTGAGGAGATCTCTGCGAGTCTGTTTCCCG 834

RESULT 4
LOCUS      A76569                      2790 bp      DNA      linear      PAT 19-OCT-1999
DEFINITION Sequence 1 from Patent WO9317110.
ACCESSION  A76569
VERSION     A76569.1  GI:6088480
KEYWORDS
SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 2790)
AUTHORS     Rodgers, B.C., and Parker, D.
TITLE       A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE
JOURNAL     Patent: WO 9317110-A 1 02-SEP-1993;
            WELLCOME FOUND (GB)
FEATURES
            Location/Qualifiers
            1..2790
            /organism="unidentified"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32644"
            1..2790
            /note="unnamed protein product; Fusion of non-structural
            and structural protein sequences of PT-NANBH"

ORIGIN
Query Match      96.5%; Score 804.8; DB 6; Length 2790;
Best Local Similarity 99.5%; Pred. No. 7e-186;
Matches 828; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY      4 AAAACCAACGTAACCAACCTCCGCCACAGAGAGTCAAGTTCCTCGGCGGTGCTCAG 63
DB      1876 AAAACCAACGTAACCAACCTCCGCCACAGAGAGTCAAGTTCCTCGGCGGTGCTCAG 1935
QY      64 ATGTTGTGTGAGTTTACCTGTTCCGCGCAGAGGCGCCAGGTTGGGTGTGCGCGACT 123
DB      1936 ATGTTGTGTGAGTTTACCTGTTCCGCGCAGAGGCGCCAGGTTGGGTGTGCGCGACT 1995
QY      124 AAGAAACTTTCCGAGGCGGTGCGCACTGCTGGAAGGCGCAACTATCCCAAGGCTGCC 183
DB      1996 AAGAAACTTTCCGAGGCGGTGCGCACTGCTGGAAGGCGCAACTATCCCAAGGCTGCC 2055
QY      184 CAGCCGAGAGGCGAGGCTGAGGCTCAGCCGAGGATACCTTGGCCCTCTATGGAACGAG 243
DB      184 CAGCCGAGAGGCGAGGCTGAGGCTCAGCCGAGGATACCTTGGCCCTCTATGGAACGAG 243

```

Db	2056	CAGCCCGAGGGCAGGGGCTGTGGGCTCAGCCCGGGATACCTTGGCCCTCTAATGGACAAGAG	2115
QY	244	GGCATTGGGGTGGGAGAGATTGGCTCTCTGTACACCCCGTGGCTCCCGGCTCAATTGGGGGCCCC	303
Db	2116	GGCATTGGGGTGGGAGAGATTGGCTCTCTGTACACCCCGTGGCTCCCGGCTCAATTGGGGGCCCC	2175
QY	304	ACTGACCCTCCGGCGTAAAGTGCCTGAATTTGGGGTAAAGTCATCGAATACCTTCACATGCGGC	363
Db	2176	ACTGACCCTCCGGCGTAAAGTGCCTGAATTTGGGGTAAAGTCATCGAATACCTTCACATGCGGC	2235
QY	364	TTCCCGCACTCTCATGSGGGTCAATTCCGCTCTGTGGCGCTCCCTTA-GGGGCGCTGCCAG	422
Db	2236	TTCCCGCAG-CTCATGSGGGTCAATTCCGCTCTGTGGCGCTCCCTTAAGGGGCGCTGCCAG	2294
QY	423	GGCGCTGGCGGATGGCGTCCGGGTTCTGGAGAGAGCGCGTGAACATATGCAACAGGGAATTT	482
Db	2295	GGCGCTGGCGGATGGCGTCCGGGTTCTGGAGAGAGCGCGTGAACATATGCAACAGGGAATTT	2354
QY	483	ACCGGGTGTCTTTCCTATCTTCTCTTGGCTTTGGCTGTGACCTTTGACCATTTCCAGC	542
Db	2355	ACCGGGTGTCTTTCCTATCTTCTCTTGGCTTTGGCTGTGACCTTTGACCATTTCCAGC	2414
QY	543	TTCCGCTTAATAGTGGCGCAAGTGTCCGGATCTAACATGTCAAGACGATTCCTCAA	602
Db	2415	TTCCGCTTAATAGTGGCGCAAGTGTCCGGATCTAACATGTCAAGACGATTCCTCAA	2474
QY	603	CTCAAGATCGTGTACGAGACAGGGGACATATATATGACACACCCCGGTTGTGGCCTG	662
Db	2475	CTCAAGATCGTGTGTACGAGACAGGGGACATATATATGACACACCCCGGTTGTGGCCTG	2534
QY	663	TGTTCGGAGAGGTAAATTCCTCCCGCTGTGGGTAGCGCTCACTCCACGCTCGGGCCAA	722
Db	2535	TGTTCGGAGAGGTAAATTCCTCCCGCTGTGGGTAGCGCTCACTCCACGCTCGGGCCAA	2594
QY	723	GGACGCCAGCATCCCACTGCGACAAATACGACGCCACGTGATTTGCTGTTGGGGCGGC	782
Db	2595	GGACGCCAGCATCCCACTGCGACAAATACGACGCCACGTGATTTGCTGTTGGGGCGGC	2654
QY	783	TGCGCTTCGTCGCGCTAATGTACGTGGGGAGATCTTCGCGAATCTGTTTTCCCG	834
Db	2655	TGCGCTTCGTCGCGCTAATGTACGTGGGGAGATCTTCGCGAATCTGTTTTCCCG	2706

RESULT	5			
A76594				
LOCUS				
DEFINITION	A76594	3372 bp	DNA	linear
ACCESSION	Sequence 26 from Patent WO9317110.			PAT 19-OCT-1999
VERSION	A76594			
KEYWORDS	A76594.1 GI:6088505			
SOURCE	.			
ORGANISM	unidentified unidentified			
REFERENCE	unclassified.			
AUTHORS	1 (bases 1 to 3372)			
TITLE	Rodgers,B.C. and Parker,D.			
JOURNAL	A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE Patent: WO 9317110-A 26 02-SEP-1993; WELLCOME FOUND (GB)			
FEATURES				
source	Location/Qualifiers			
	1..3372			
	/organism="unidentified"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:32644"			
CDS	1..>3372			
	/note="unnamed protein product; Encodes a fusion protein which can be expressed in insect cells infected with the baculovirus BHC-28"			
	/codon_start=1			
	/protein_id="CAB58632.1"			
	/db_xref="GI:6088506"			
	/translation="MPDYSPRPTGDPDPSLSAFTEVDGVRLHRYAPACKPLIREV TFQGLNQYLVGSQLPCPEPBDVAVLITMLDPSHTIAETAKRIKLASGPSPLASSA SQSGPSSKATYITQNDPPADLLLEANLWMHEWGGITRVESSENKVILIDSFDPLEA			

EEBERVSVAEILIRKSKKPEPMAMPARNDYPPLESKADVDYVPVHGGCELPY  
 KTPPIPEPRKRTVVLTESTVSALAEIATKADYSGSASVDSGTATVPADGSDOG  
 AGSDVSSYMSMLPEEGEPDPLSDSGSWTSEAEADYCSWCTWICALITPCQA  
 BESEKIINLMSNLLHHNVVYVITRSASQOROKYTPRLOJLIDHGXOUIEMKAK  
 ASTVKAKLISVEACKLITPHSKSKRFGAKOVRNLSKAIHHISWEDLIEDTET  
 PAVTAKKLSVEACKLITPHSKSKRFGAKOVRNLSKAIHHISWEDLIEDTET  
 PLITITMAKNEVFCVQBERGKRKAPLIVFDIGVCEKENDIYVSEIYQCDLAP  
 YGQYSPQORVEFLVNAUMSKKTPMGPAVDITFDSVITENDIYVSEIYQCDLAP  
 ARCAISRLTERLIYIGPLDNLNSKNGARVYLARATPGSTVTPHPIEALASNTGII  
 DECHSDSTLSLIGITGVLPDQAEATAGARVYLARATPGSTVTPHPIEALASNTGII  
 IPEFGKAIPIETIIKGRHLLIFCHSKKQDELAALVGLGAINAVYSGLDVSVIPASD  
 VVVVUADLMTGTGDEDSVIDCNCVTQOVBDELPTFIETTVTPOBDAVSORRG  
 RARGSGRGYRFAVAARVKKKKKTKRNTNINPOOVRPGGOIYGVVYLPERPRJG  
 RARTKTSBSQPRGRORPIKARQOPBGMAARVYVPGVNGNMGMAWGLNIPDSGR  
 PMSGPNDBPRRSPNLGKVIDITLTCGRADLMTGYVLVAGLIGARALALAHVRLTEGVC  
 NYATONLPGCSSEPIFLALISCTIIPASAEVYNVAGIHYNTVDLSSAFCYMTAM  
 MHTPGVCPCVBNESNCWALPITLAKDAGSLPTATIRRHVDLGVGAALFCSAMTGM  
 DLGGSVPEPEQLSAGRIGSPGTRQEPKTHISLOGNP

Query Match	96.54%	Score 804.8	DB 6	Length 3372	
Best Local Similarity	99.54%	Pred. No. 7e-186			
Matches 82;	Conservative 0	Mismatches 2	Indels 2	Gaps 2	
QY	4	AAAAACAAACGTAAACAACCAACTCCGCCACAGACGTCAGGTTCCGGGCGGTGCTCAG	63		
Db	2458	AAAACCAAACGTAAACAACCAACTCCGCCACAGACGTCAGGTTCCGGGCGGTGCTCAG	2517		
QY	64	ATCGTTGTGTGAAGTTAACTGTGTGCGCGGACAGGGGCCCCAAGGTTGGGTGTGCCGCACT	123		
Db	2518	ATCGTTGTGTGAAGTTAACTGTGTGCGCGGACAGGGGCCCCAAGGTTGGGTGTGCCGCACT	2577		
QY	124	AGGAAGACTTCGAGAGGGTGTGCAACTCCGTGGAAGGACAACTATCCCAAGGCTCGC	183		
Db	2578	AGGAAGACTTCGAGAGGGTGTGCAACTCCGTGGAAGGACAACTATCCCAAGGCTCGC	2637		
QY	184	CAGCCCGAGGGCAAGGCTGTGGCTCAGCCCGGGTACCTTTGGCCCTCTATGTGCAACGAG	243		
Db	2638	CAGCCCGAGGGCAAGGCTGTGGCTCAGCCCGGGTACCTTTGGCCCTCTATGTGCAACGAG	2697		
QY	244	GGCATGAGGGTGGGACAGATGGCTCTGTGTCACCCCGTGGCTCCGGCCCTAAGTTGGGCCCC	303		
Db	2698	GGCATGAGGGTGGGACAGATGGCTCTGTGTCACCCCGTGGCTCCGGCCCTAAGTTGGGCCCC	2757		
QY	304	ACTGACCCCCGGGGTGAAGTCGTAAATTTGGGTAAAGTACATGATACCTCAATGCGGC	363		
Db	2758	ACTGACCCCCGGGGTGAAGTCGTAAATTTGGGTAAAGTACATGATACCTCAATGCGGC	2817		
QY	364	TTTGCCGCACTCTATGAGGGGTACATTCCTCCGCTCGTCGGGCGCTCCCTTA-GGGGCGCTGCGAG	422		
Db	2818	TTTGCCGCACTCTATGAGGGGTACATTCCTCCGCTCGTCGGGCGCTCCCTTAAGGGGGCGCTGCGAG	2876		
QY	423	GGCCCTGGCGCATGGGCGTCCGGGTTCTGAGAGACGGCGTGAATATATGCAACAGGGAATTT	482		
Db	2877	GGCCCTGGCGCATGGGCGTCCGGGTTCTGAGAGACGGCGTGAATATATGCAACAGGGAATTT	2936		
QY	483	ACCCGGTTGCTCTTTCTCTATCTTCTCTCTTGGCTGTGCTGTGTTGACATTTCCAGC	542		
Db	2937	ACCCGGTTGCTCTTTCTCTCTCTTCTCTCTTGGCTGTGCTGTGTTGACATTTCCAGC	2996		
QY	543	TTCCGCTTAAAGTGCAGACGTCGCCGGAATCTACATGTACGAACGATTTGCTCCAA	602		
Db	2997	TTCCGCTTAAAGTGCAGACGTCGCCGGAATCTACATGTACGAACGATTTGCTCCAA	3056		
QY	603	CTCAAGCATGTGTATGAGACAGACGGAATGATCATGACACCCCGGGGTGTGCCCTG	662		
Db	3057	CTCAAGCATGTGTATGAGACAGACGGAATGATCATGACACCCCGGGGTGTGTGCCCTG	3116		
QY	663	TGTCTGGGAGGGTAATTTCTCCCGCTGTGGGTAGCGCTCACTCCACGCTGCGGCGAA	722		
Db	3117	TGTCTGGGAGGGTAATTTCTCCCGCTGTGGGTAGCGCTCACTCCACGCTGCGGCGAA	3176		
QY	723	GGACGCCAAGCATCCCACTGCGCAATAGACGCCACGTCGATTTGCTGTTGGGGCGGC	782		

Db 3177 GGCAGCCAGCATCCCACTGCGCAACAATACGACCGCAAGTCGATTCCTGTTGGGGCGGC 3236

QY 783 TGGCTTCCTGCTCGCTATGATGAGGGGAGTCTGCGGATCTGTTTCCG 834

Db 3237 TGGCTTCCTGCTCGCTATGATGAGGGGAGTCTGCGGATCTGTTTCCG 3288

RESULT 6

A28155

LOCUS 2116 bp DNA linear PAT 07-JUN-1995

DEFINITION PT-NANBH mRNA fragment from patent GB2239245.

ACCESSION A28155

VERSION A28155.1 GI:1248638

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1.2116

Post-transfusal non-A non-B hepatitis viral polypeptides

Patent: GB 2239245-A 21 26-JUN-1991;

Location/Qualifiers

1..2116

codon\_start=1

/transl\_table=11

/product="PT-NANBH polypeptide N-terminus"

/protein\_id="CAA01931.1"

/db\_xref="GI:1248638"

308..>2116

contig formed by cDNA clones"

ORIGIN

Query Match 96.2%; Score 802.6; DB 6; Length 2116;

Best Local Similarity 99.3%; Pred. No. 2.4e-185;

Matches 827; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 AGAAAAACCAACGTAACCAACCTCCGCCACAGAGGTCCCGGGCGGTGGT 60

Db 333 AGAAAAACCAACGTAACCAACCTCCGCCACAGAGGTCCCGGGCGGTGGT 391

QY 61 CAGATCGTTGTGAGATTAACCTGTTCCGCGGAGGGGCGGAGTTGGGTGCGCGG 120

Db 392 CAGATCGTTGTGAGATTAACCTGTTCCGCGGAGGGGCGGAGTTGGGTGCGCGG 451

QY 121 ACTAGAGAGACTTCCGAGCGGTGCAACTCGTGAAGGAGCAACCTATCCCAAGGCT 180

Db 452 ACTAGAGAGACTTCCGAGCGGTGCAACTCGTGAAGGAGCAACCTATCCCAAGGCT 511

QY 181 CCGCAGCCGAGAGGCGGCGGTGCGGAGCGGCGGAGTACCTTGGCCCTCTATGGCAAC 240

Db 512 CCGCAGCCGAGAGGCGGCGGTGCGGAGCGGCGGAGTACCTTGGCCCTCTATGGCAAC 571

QY 241 GAGGCGATGGGTTGGGCGAGATGCTCTGTCAACCGGTGAGTCCCGGCGCTAGTGGGG 300

Db 572 GAGGCGATGGGTTGGGCGAGATGCTCTGTCAACCGGTGAGTCCCGGCGCTAGTGGGG 631

QY 301 CCGACTGACCCCGCGGAGTGGCTGCGTAATTGGGTAAGTCATGATACCTCATATGC 360

Db 632 CCGACTGACCCCGCGGAGTGGCTGCGTAATTGGGTAAGTCATGATACCTCATATGC 691

QY 361 GCGTTCGAGCACTCTCATGGGGATACATTCGCTCGTGGCGCTCCCTTA-GGGGCGCTGC 419

Db 692 GCGTTCGAGCACTCTCATGGGGATACATTCGCTCGTGGCGCTCCCTTAAGGGGCGCTGC 750

QY 420 CAGGGCCCTGCGCATGAGCGCTCGGTTCTGAGAGAGCGCGTGAATATGCAACAGGAA 479

Db 751 CAGGGCCCTGCGCATGAGCGCTCGGTTCTGAGAGAGCGCGTGAATATGCAACAGGAA 810

QY 480 TTTTACCCGCTGCTCTTCTCTATCTCTCTCTTGGCTTGGCTCTCTGTTTGAACATTC 539

Db 811 TTTTACCCGCTGCTCTTCTCTATCTCTCTCTTGGCTTGGCTCTCTGTTTGAACATTC 870

QY 540 AGCTTCGCTTGAAGTGGAGCGAAGTGTCCGGAGCTCAATGTACAGCAAGATTCCTC 599

Db 871 AGCTTCGCTTGAAGTGGAGCGAAGTGTCCGGAGCTCAATGTACAGCAAGATTCCTC 930

QY 600 CAATCAGCATCTGTATGAGAGAGAGCATGATCATGACACCCCGGGGTGTGCG 659

Db 931 CAATCAGCATCTGTATGAGAGAGAGCATGATCATGACACCCCGGGGTGTGCG 990

QY 660 CTGTGCTCGGAGAGGTTATTTCTCCGCTGCTGGTGAAGGCTCCTCCACGCTCGCGGC 719

Db 991 CTGTGCTCGGAGAGGTTATTTCTCCGCTGCTGGTGAAGGCTCCTCCACGCTCGCGGC 1050

QY 720 CAGAGCGCGAGATCCCACTGCGCAACAATACGAGCGCAAGTCGATTTGCTGGGGG 779

Db 1051 CAGAGCGCGAGATCCCACTGCGCAACAATACGAGCGCAAGTCGATTTGCTGGGGG 1110

QY 780 GCGTCCCTTCTGCTCGCTATGATGATGAGGGAGTCTCTGCGATCTGTTTCC 832

Db 1111 GCGTCCCTTCTGCTCGCTATGATGATGAGGGAGTCTCTGCGATCTGTTTCC 1163

RESULT 7

A32202

LOCUS 2116 bp DNA linear PAT 10-DEC-1996

DEFINITION NANBH PT polypeptide fragment.

ACCESSION A32202

VERSION A32202.1 GI:1926562

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

FEATURES

source

1.2116

Non-A, non-B hepatitis virus

Non-A, non-B hepatitis virus

Viruses; unclassified viruses.

1 (bases 1 to 2116)

Patent: FR 2655990-A 21 21-JUN-1991;

Location/Qualifiers

1..2116

codon\_start=1

/organism="Non-A, non-B hepatitis virus"

/mol\_type="unassigned DNA"

/db\_xref="taxon:12440"

308..>2116

/product="partial NANBH-PT polypeptide"

/protein\_id="CAA02094.1"

ORIGIN

Query Match 96.2%; Score 802.6; DB 6; Length 2116;

Best Local Similarity 99.3%; Pred. No. 2.4e-185;

Matches 827; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 AGAAAAACCAACGTAACACCAACCTCCGCCACAGAGCTCAGGTTCCGGGCGGTGCT 60  
DB 332 AGAAAAACCAACGTAACACCAACCTCCGCCACAGAGCTCAGGTTCCGGGCGGTGCT 391  
QY 61 CAGATCGTTGGTGAGGTTTACCTGTTGCGCGCAGAGGGGCCCAAGTTGGTGTGCGCG 120  
DB 392 CAGATCGTTGGTGAGGTTTACCTGTTGCGCGCAGAGGGGCCCAAGTTGGTGTGCGCG 451  
QY 121 ACTAGGAAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGACAACCTATCCCAAGCT 180  
DB 452 ACTAGGAAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGACAACCTATCCCAAGCT 511  
QY 181 CGCCAGCCCGAGAGGCGAGGCTTGAGCTCAGCCCGGATACCTTGAGCCCTCTATGCAAC 240  
DB 512 CGCCAGCCCGAGAGGCGAGGCTTGAGCTCAGCCCGGATACCTTGAGCCCTCTATGCAAC 571  
QY 241 GAGGCGATGAGGTTGGGCGAGAGTGGCTCCGTGACCCCGGTGCTCCCGGCTATGTTGGGCG 300  
DB 572 GAGGCGATGAGGTTGGGCGAGAGTGGCTCCGTGACCCCGGTGCTCCCGGCTATGTTGGGCG 631  
QY 301 CCCACTGAACCCCGGCGGTAGGTGCGGTAAATTTGGTAAAGTCATGATACCTTCACATGC 360  
DB 632 CCCACTGAACCCCGGCGGTAGGTGCGGTAAATTTGGTAAAGTCATGATACCTTCACATGC 691  
QY 361 GAGTTGCGCGACTCATGAGGTTGATTCGCTGCTGCGCGCTCCCTTA-6GAGCGCTGC 419  
DB 692 GAGTTGCGCGACTCATGAGGTTGATTCGCTGCTGCGCGCTCCCTTA-6GAGCGCGCTGC 750  
QY 420 CAGGCGCTGAGGCGAGTGGCGTCCGGGTTCTGAGAGACCGCGGTAAGTATGCAACAGGAA 479  
DB 751 CAGGCGCTGAGGCGAGTGGCGTCCGGGTTCTGAGAGACCGCGGTAAGTATGCAACAGGAA 810  
QY 480 TTTAACCCGTTGCTTTCTCTATCTTCTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539  
DB 811 TTTAACCCGTTGCTTTCTCTATCTTCTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870  
QY 540 AGCTTCGCGCTTATGAAGTGCAGACGTGTCCGGGATCTAACATGTCACGAACGATTGCTC 599  
DB 871 AGCTTCGCGCTTATGAAGTGCAGACGTGTCCGGGATCTAACATGTCACGAACGATTGCTC 930  
QY 600 CAACCTCAAGCATCGTGTACGAGACGCGGACATGATGCAACCCCGGTTGTGCTC 659  
DB 931 CAACCTCAAGCATCGTGTACGAGACGCGGACATGATGCAACCCCGGTTGTGCTC 990  
QY 660 CTGTGTCCGGAGGAGTAAATTTCTCCGCTGCTGAGTGGCTCACTCCACGCTCGCGC 719  
DB 991 CTGTGTCCGGAGGAGTAAATTTCTCCGCTGCTGAGTGGCTCACTCCACGCTCGCGC 1050  
QY 720 CAAGACGCGAGCATCCCACTGCGACAATACGAGCGCACTGCAATTTGCTGTTGGGCG 779  
DB 1051 CAAGACGCGAGCATCCCACTGCGACAATACGAGCGCACTGCAATTTGCTGTTGGGCG 1110  
QY 780 GAGTGCCTTCTGCTCCGCTATGATAGTGGGAGATCTGCGGATCTGTTTTC 832  
DB 1111 GAGTGCCTTCTGCTCCGCTATGATAGTGGGAGATCTGCGGATCTGTTTTC 1163

RESULT 8  
ARI44050 2116 bp DNA linear PAT 08-AUG-2001  
LOCUS ARI44050  
DEFINITION Sequence 21 from patent US 6210675.  
ACCESSION ARI44050  
VERSION ARI44050.1 GI:15105917  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2116)  
AUTHORS Higfield, P. Edmund., Rodgers, B. Colin., Tedder, R. Seton. and  
TITLE PT-NANB hepatitis polypeptides  
JOURNAL Patent: US 6210675-A 21 03-APR-2001;

FEATURES  
Source Location/Qualifiers  
1..2116  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 96.2%; Score 802.6; DB 6; Length 2116;  
Best Local Similarity 99.3%; Pred. No. 2.4e-185;  
Matches 827; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 AGAAAAACCAACGTAACACCAACCTCCGCCACAGAGCTCAGGTTCCGGGCGGTGCT 60  
DB 332 AGAAAAACCAACGTAACACCAACCTCCGCCACAGAGCTCAGGTTCCGGGCGGTGCT 391  
QY 61 CAGATCGTTGGTGAGGTTTACCTGTTGCGCGCAGAGGGGCCCAAGTTGGTGTGCGCG 120  
DB 392 CAGATCGTTGGTGAGGTTTACCTGTTGCGCGCAGAGGGGCCCAAGTTGGTGTGCGCG 451  
QY 121 ACTAGGAAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGACAACCTATCCCAAGCT 180  
DB 452 ACTAGGAAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGACAACCTATCCCAAGCT 511  
QY 181 CGCCAGCCCGAGAGGCGAGGCTTGAGCTCAGCCCGGATACCTTGAGCCCTCTATGCAAC 240  
DB 512 CGCCAGCCCGAGAGGCGAGGCTTGAGCTCAGCCCGGATACCTTGAGCCCTCTATGCAAC 571  
QY 241 GAGGCGATGAGGTTGGGCGAGAGTGGCTCCGTGACCCCGGTGCTCCCGGCTATGTTGGGCG 300  
DB 572 GAGGCGATGAGGTTGGGCGAGAGTGGCTCCGTGACCCCGGTGCTCCCGGCTATGTTGGGCG 631  
QY 301 CCCACTGAACCCCGGCGGTAGGTGCGGTAAATTTGGTAAAGTCATGATACCTTCACATGC 360  
DB 632 CCCACTGAACCCCGGCGGTAGGTGCGGTAAATTTGGTAAAGTCATGATACCTTCACATGC 691  
QY 361 GAGTTGCGCGACTCATGAGGTTGATTCGCTGCTGCGCGCTCCCTTA-6GAGCGCTGC 419  
DB 692 GAGTTGCGCGACTCATGAGGTTGATTCGCTGCTGCGCGCTCCCTTA-6GAGCGCGCTGC 750  
QY 420 CAGGCGCTGAGGCGAGTGGCGTCCGGGTTCTGAGAGACCGCGGTAAGTATGCAACAGGAA 479  
DB 751 CAGGCGCTGAGGCGAGTGGCGTCCGGGTTCTGAGAGACCGCGGTAAGTATGCAACAGGAA 810  
QY 480 TTTAACCCGTTGCTTTCTCTATCTTCTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539  
DB 811 TTTAACCCGTTGCTTTCTCTATCTTCTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870  
QY 540 AGCTTCGCGCTTATGAAGTGCAGACGTGTCCGGGATCTAACATGTCACGAACGATTGCTC 599  
DB 871 AGCTTCGCGCTTATGAAGTGCAGACGTGTCCGGGATCTAACATGTCACGAACGATTGCTC 930  
QY 600 CAACCTCAAGCATCGTGTACGAGACGCGGACATGATGCAACCCCGGTTGTGCTC 659  
DB 931 CAACCTCAAGCATCGTGTACGAGACGCGGACATGATGCAACCCCGGTTGTGCTC 990  
QY 660 CTGTGTCCGGAGGAGTAAATTTCTCCGCTGCTGAGTGGCTCACTCCACGCTCGCGC 719  
DB 991 CTGTGTCCGGAGGAGTAAATTTCTCCGCTGCTGAGTGGCTCACTCCACGCTCGCGC 1050  
QY 720 CAAGACGCGAGCATCCCACTGCGACAATACGAGCGCACTGCAATTTGCTGTTGGGCG 779  
DB 1051 CAAGACGCGAGCATCCCACTGCGACAATACGAGCGCACTGCAATTTGCTGTTGGGCG 1110  
QY 780 GAGTGCCTTCTGCTCCGCTATGATAGTGGGAGATCTGCGGATCTGTTTTC 832  
DB 1111 GAGTGCCTTCTGCTCCGCTATGATAGTGGGAGATCTGCGGATCTGTTTTC 1163

RESULT 9  
A76571 831 bp DNA linear PAT 19-OCT-1999  
LOCUS A76571  
DEFINITION Sequence 3 from Patent WO9317110.  
ACCESSION A76571  
VERSION A76571.1 GI:6088482



LOWIPLNVRGGDAIILLTCVAHPELIDITKLLALLGLPLMTLQAGIIRAPYFVR  
AOGILRACMLVRKAVGYVQWALMKLALVTGTYYDHLTPIDMAHGLDLAIVT  
PVFSDMETKXITMGADTAAGDVLILGVALSRGKEIFLGDALIEGQDLAATP  
AYSOOTRGLIGLITSLTRGKNVGEVGVSTANOSRIATCNVQMTVPHRGSK  
TLAQRGPIPTOMTIVDODLVWPAFPAKRSITPTCTCCSDIVLVRHADIVPHRG  
DSRGLSPRPVSYLKSSGGPLCPSHVAGIIPRAACSTDRVAALVFPVPESEET  
TRSPVTNDSPPAVPQTFQVNAHHAIPGSGKSTPVPAVAQKAVLUNBSVATL  
SFGAYMSKAGVDPISIRGTQRTITGAPITYSTYGTFLADGCSGAADIIICDCHS  
TDSITLIGITLDOAETAGARLVLTAPGVSATVPHPNIEVALPSTGDIVYVKA  
ALPIEIRGGGHLIFCHSKKCDLAALOSLGVNAVAVYRGDLVSVIPSGDIPVVA  
TDALMTGTGDFDSVIDCTCTQTVVDRLSDPTIETITTVQDVAESSORGRTRG  
KRGIYRFPVGPSPGMDSSVLCEDYDAGCAWELTPAETSVLRALANTPLGVPQC  
DHLFEWSEVFTGLTHIDAFLOQKQADPBYLAQVTCARQKAVLPBMDQMKC  
LIRKPTLHGPPLRLGAVQNEVTLHPRTYIMACADLEVTWVTVGVGLA  
ALLAYCTTGSVIVIGRIILSGPAIIPREVLRYRBPDMESCAHLPYIBGOMLAE  
OPEKALILQOTKQEAAPAVESKQALAEFANMMNNTISGVYIAGISTIPGN  
PAISLMAFTASITSPITTOHTLNLILGAVQAOLAPPSAAAVAGAGIAGAVCAA  
GIGKVLVDILAGYAGVAGALVAFKVMGEMSTEDLVNLLPALISPALVAVGCAI  
KLRRHVGBEGAVQMMNRLIAPSRGNHVSPTHYVESDPAARVQIILSLITLQLK  
RLHQMISDCSTPSCGSMRLDWDWICTVLDTFKMLQSLPLRGVFPFCORGK  
GVMRGDMQOTPCGAQITGHVKNKSMRIVGPTCSMTMGTPTPINKVYTGCTPSP  
APNYSRALMRVAEEVETRVGDPHVYTGMTDNDVCKPCCPVAREPTEVDGVLHR  
YAPACRPLIRDEVTQVGLNOYLVSQLPCEBEPVYVITSMITDPSHITPAARL  
AKGSPSLASSASQSLAPSLATCTTHDSFDADLIEANLIMRQEMGNITRVSEN  
KVLILSDPDLAEDEBERISVPAKILKRSFPPAMPIMAPDNPPLIESMDPDY  
VPPVHGGPLPTTRAPPIPPPKKRTVLTSTVSALAEATKTFGSGSASVDSGT  
ATAPDOPSDGDDTGSVYESMPLEBEPDPLCSMPLTEEBESBDDVCSMS  
YTMGALITPCAAEESKLPINALNSLIRHNHMYATVTSRASOROKKXVTPRLQVLD  
DHYRVLKEMRAKASTVYKARLISBEACULPTPHSKSKFPGAGADVNLISKALNHI  
NSWEDLJEDTETPIDITIMAKNEVFCQPEKGRKPAFLVFDLGVAVCKMALYD  
VSTLPHAVMSSYGFQYSPQGRVLYLVKAMSSKNPQFAEDTCEFTSTVLEDIRV  
BESITQCCDLAPEARQAIRSLTERLYIGGLPNSGQNCGYRCAAGLTTSCENTL  
TCYLKASACRAKLODCMTLMDGDLVLYICSAQTOEABSLARVETEMTRESAPG  
DPBOEYVLELITSCSNVSVAHDSGRVYVLTDPPTTLARAAWETAARHPVMSLT  
GNITMYATLWARMYLMTHFESILIAQOBLEALDCQIYGACYSIEPDIPIORLH  
GISAPLSHSYSPGELNRYVASCIRKLGVPPLRWRRARSVRKLSIRGRATCTKYL  
FNMAVTRKLLKPIPAASRLDISGMFVAVGYSQDLYHLSIPARPMFWMLLSVGV  
GIYLPNR"

ORIGIN

Query Match	89.7%;	Score 748.2;	DB 14;	Length 9533;
Best Local Similarity	95.2%;	Pred. No. 4.6e-172;		
Matches 793;	Conservative 0;	Mismatches 38;	Indels 2;	Gaps 2;
QY	1	AGAAAAACCAACGTAACCAACCACTCCGCCACAGAGAGTTCAGGTTCCCGGCGGTGT	60	
DB	365	AGAAAAACCAACGTAACCAACCACTCCGCCACAGAGAGTTCAGGTTCCCGGCGGTGT	424	
QY	61	CAGATCGTTGGTGAAGTTACTGTTCGCGCAGAGGGCCCGAGTTGGGTGCGCGG	120	
DB	425	CAGATCGTTGGTGAAGTTACTGTTCGCGCAGAGGGCCCGAGTTGGGTGCGCGG	484	
QY	121	ACTAGAGAACTTCCGAGCGGTGCGCACTCGTGAAGCGCAACCTATCCCAAGGCT	180	
DB	485	ACTAGAGAACTTCCGAGCGGTGCGCACTCGTGAAGCGCAACCTATCCCAAGGCT	544	
QY	181	CGCCAGCCCGAGAGCGAGGCTGAGCGCTAGCCGGGTACCTTGGGCCCTTATGGCAAC	240	
DB	545	CGCCAGCCCGAGAGGTAAGGCTGAGCGCTAGCCGGGTACCTTGGGCCCTTATGGCAAC	604	
QY	241	GAGGGCATGGGGTGGGACAGATGCTCTGTCAACCCCGTGGCTCCCGGCTTAGTTGGGGC	300	
DB	605	GAGGGCATGGGGTGGGACAGATGCTCTGTCAACCCCGTGGCTCCCGGCTTAGTTGGGGC	664	
QY	301	CCCACTGACCCCGGCGGTAGGTTCGCTAATTTGGGTAAGTCATGATACCTTCACATGC	360	
DB	665	CCCACTGACCCCGGCGGTAGGTTCGCTAATTTGGGTAAGTCATGATACCTTCACATGC	724	
QY	361	GAGCTTGGCGGACCTTCATGAGGGTACATTCGCGCTCGTGGCGCTCCCTTA-GGGGGCGTGC	419	
DB	725	GAGCTTGGCGGAC-CTTCATGAGGGTACATTCGCGCTCGTGGCGCTCCCTTA-GGGGGCGATGC	783	
QY	420	CAGGGCCCTGGCGCATGGCGCTCGGGGTTCTGAGAGGACGGCGTGAACCTATGCAACAGGGAA	479	

DB	784	CAGGGCCCTGGCGCATGGCGCTCGGGGTTCTGAGAGGACGGCGTGAATTATGCAACAGGGAA	843	
QY	480	TTTACCAGGTGCTCTTCTCATATCTCTCTGCGCTTTCGCTGCTTTGACCATTC	539	
DB	844	TTTGCCTGGGCTCTTCTCTCATATCTCTCTGCGCTTTCGCTGCTTTGACCATTC	903	
QY	540	AGCTTCCCTTATGAAAGCGCCAAAGTGTCCGGGATCTACATGTCAAGAAAGATTGCTC	599	
DB	904	AGCTTCCCTTATGAAAGCGCCAAAGTGTCCGGGATCTACATGTCAAGAAAGATTGCTC	963	
QY	600	CAACTCAAGCATCTGTGACAGAGACAGCGGACATGATCAACACCCCGGGGTGTGCTC	659	
DB	964	CAACTCAAGCATCTGTGTGAGACAGCGGACATGATCAACACCCCGGGGTGTGCTC	1023	
QY	660	CTGTGCTCGGAGAGGTATATCTCCCGGTGCTGGGTACGGCTCACTCCACGCTGCGGC	719	
DB	1024	CTGTGCTCGGAGAGGTATATCTCCCGGTGCTGGGTACGGCTCACTCCACGCTGCGGC	1083	
QY	720	CAAGAGCCCGACATCCCGCACTGCGACATAAGACGCGACGTCGATTGCTGTGGGCG	779	
DB	1084	CAGAAATGAGACGCTCCCACTACGCGCATAGACGCAACGTCGATTGCTGTGGGCG	1143	
QY	780	GAGTCCCTTCTGCTCGCTATGATGACGTGGGGATCTTGGCGATCTTTTCC	832	
DB	1144	GAGTCCCTTCTGCTCGCTATGATGACGTGGGGATCTTGGCGATCTTTTCC	1196	

RESULT 11

HPCNSISPF 1595 bp ss-RNA linear VRL 02-AUG-1993  
LOCUS Hepatitis C virus core protein (NS1) gene.  
DEFINITION M74809  
ACCESSION M74809.1 GI:329787  
VERSION  
KEYWORDS NS1 protein; core protein; envelope-associated protein.  
SOURCE Hepatitis C virus  
ORGANISM Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1 (bases 1 to 1595)  
AUTHORS Cha,T.A., Beall,E., Irvine,B., Kolberg,J., Chlen,D., Kuo,G. and Urdea,M.S.  
TITLE At least five related, but distinct, hepatitis C viral genotypes exist

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (15), 7144-7148 (1992)  
MEDLINE 92357788  
PUBMED 1323128  
COMMENT Original source text: Hepatitis C virus (individual isolate 115)  
CDNA to genomic RNA.  
FEATURES  
source  
1..1595  
Location/Qualifiers  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/isolate="115"  
/db\_xref="taxon:11103"

ORIGIN

Query Match	89.5%;	Score 746.6;	DB 14;	Length 1595;
Best Local Similarity	95.1%;	Pred. No. 1.2e-171;		
Matches 792;	Conservative 0;	Mismatches 39;	Indels 2;	Gaps 2;
QY	1	AGAAAAACCAACGTAACCAACCACTCCGCCACAGAGAGTTCAGGTTCCCGGCGGTGT	60	
DB	266	AGAAAAACCAACGTAACCAACCACTCCGCCACAGAGAGTTCAGGTTCCCGGCGGTGT	325	
QY	61	CAGATCGTTGGTGAAGTTACTGTTCGCGCAGAGGGCCCGAGTTGGGTGCGCGG	120	
DB	326	CAGATCGTTGGTGAAGTTACTGTTCGCGCAGAGGGCCCGAGTTGGGTGCGCGG	385	
QY	121	ACTAGAGAACTTCCGAGCGGTGCGCACTCGTGAAGCGCAACCTATCCCAAGGCT	180	
DB	386	ACTAGAGAACTTCCGAGCGGTGCGCACTCGTGAAGCGCAACCTATCCCAAGGCT	445	

```

RESULT 12
AF207768          LOCUS          9379 bp      RNA      linear      VRL 27-APR-2000
DEFINITION      Hepatitis C virus strain MD27 complete genome.
ACCESSION      AF207768
VERSION        AF207768.1  GI:7650253
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS      Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Marumo,F. and Sato,C.
TITLE        Characteristics of hepatitis C viral genome associated with disease progression
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 9379)
AUTHORS      Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Izumi,N. and Sato,C.
TITLE        Direct Submission
JOURNAL      Submitted (23-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan
FEATURES
SOURCE      1..9379
              /organism="Hepatitis C virus"
              /mol_type="genomic RNA"

```

CD5

```

/strain="MD21"
/db xref="taxon:11103"
/ncfe="isolated from a patient with liver cirrhosis and
hepatocellular carcinoma (HCC-7)"
330..9362
/codon_start=1
/product="polyprotein"
/strain_id="AF65958.1"
/db xref="GI:7650254"
/translation="
MSTNPKPQRKTKANTRRPDDVKFPGGGQIVGGVYLLPRRGRPL
GVATRTTSRSQRRGRKQIPKARREGGAAMQPGVWPLVGGEGAGMWLLSPGG
SRSMGCTDPRRSRNKQVIDLTLCFADLMGQIPLVGAPLAKALAGHVLGVE
GVNATLPGCSFISLLALSLCLTPASAEYRANRNGVYHTNDCSNLTVSEAD
MNTTGPCVCFREGENSRRCWALPTLARNSIPVTRRHVDLTGGAAPCSAAY
VGLDGSVPIVLSQFTTSPRRYETVODCNCSLYGHTGTHMMMDMMNSPTALV
SOLRLTQAVMDVAGAHMGLAGLITSNVGMKAKILYMLPFGADGSTRVWGGA
AQRHGFSAFSPSPGAQKIQIKTNGSWHINRPLNDCSLNLTGALDETRVWGGA
GCEBRASCRRIDKFAQGMGPIITHVNPICPMHVAPOGCIYPAEVCQPVY
FTSPVVGTTDSGAPTYMNGENETDVLILNTRPPLGMPFGTMWNSGFTTCCA
PNCIGVGANNLTICPTDCRKHPEATYVTCGSGPMLTPCWDYPRLMYPCVFA
TVQGVGNGVGHRLKACNMHMGECQLEDDBRSLSLTSTTEWMLPGFTLL
PALSTGILHLOINIVQYIYGVGSANVSIVIRNEYLLFLILADRCVACAMILL
IAQEALENLVLTNAASVAGHGISLFLVFCAMYTKEGLFGAAYALGVPLILL
LLALPRAYAMDRKEMASCGGVFGILATLISPHKVLPLAKILMWQYFTEAHEH
LOIVRPLNCRKGAQVYVOMALMKLAALGTGVYVHLPRHMAAGRLDAVFAE
AOGILYVYLVKVAAGHYVOMALMKLAALGTGVYVHLPRHMAAGRLDAVFAE
PVVPSDEKETILTMGADTLAAGDIIIGLPSARSGSLILGPDRPDEQGRRLATP
AYSQRTGLIGCITSTTGRDKNOVEBOVYVSTRNOSLILGPDRPDEQGRRLATP
TLGPKQPIOTMYTNVDQDLVGWQAPPGASMTPECTGSSDLYLVTHADVTPRRRG
DSRSLSPPRIISYLKSSSGGPLLCPSGHVGVIYFRAVCTRGVAKADVPVESMETT
MRSVFLPNDNSPPVAPQTFQVAHLMHAGTGSKSTPAAVAAQGYKVLVNPVAAATL
GFGVMSKAGVBNDRINRTGRTITTGAPITYSTYKGLADGSCGAGVILIMCECHS
TDSYSLIGITGVLDQAEETAGRLAVLATATPPGSVTVPHNPHEVALSNGEIPFYGG
AIPIETIKGSRHLIFCHSKKCCDEIAKISALGINAAYYRGDVSIVPSNGVYVAA
DLAMTGTGDPDSVIDCNTCTOTVDPSLDPTFTTETTTVPDVAVSQRGRGTGCG
RGLRFPVTPGERSPSGMDSVLCCEYDACCAMELTPAETSVRLRYLNTGPLYVCC
DLTFEFSVETGLTHIDAHPLSQTKOAGDNFYVLAQVYACAAQALPSPMDGMKAC
LHTRKPLHGPTPLLYRLGAVNOMELTLTHEITIKYIMCADAEVSTVWLVGGVLA
ALAAVCTGTSVAVGRILISGRKAVIPDEBVLFREDEMECHSLPIYEOGQVLLP
OFOKALGILTOATKOTKGAEMAAPVESKRWLEGGWAGMNNFTSGIOYGLSTLPNG
PALASMAFTASITSPLTTOHTLIFNLIGMVAQALPAPBAASFPVAGIAGAAGVSI
GLGVLDVILIAGVAGVAGALVAFKVSGEIPSADEMDMLPALSLGALGVAGVSI
ILRHHVPGEGAQVOMNRLLAFASRGNHSPETHAVPENNDAARVTOILSLTITVQLA
RLHMDINEDCTSPGCSMLRDVMDICTVLTDFETMLOSLPLPVPYFSPSCGTYSK
GWRGDMVOTCPCGAQITGVHNGKSMRYLGPASACNTHMGPIINAYATGTPSPYK
APNYSRLMFWAAVEEYVTRVGDVFNHYVTMTDNNKPCQOVPAREPTELDGRLRH
YAPCRPLREBETTFQYGLAOYPGSGLPCEBPEDVTVYLSMLDPRHITRABEARBL
ARQSPRLASSASQSLAPSLKATCTTCHDPSDDLLLEAMLMQENGTNTRVESEN
KNVILDFDPRLRAEEDEREVSVAELIRKSRPPAPLIPARPNPPLLESAGSDPGD
PPVHVHSCP.LPTKAPPIPPRRKRTVLTESYVSLALBETLTPSGSLSWADPGV
ATAPDQDLDGDSVESISWMPLEGEGRGDDLDGSGSYTYSBEAGEVDVCCSMKS
YTMGALITPCAAEESKLPITNALSILRHHNVYLTNDSASGROKKVFPDRLOVLE
DHRDVLKEMKAKASTYKAKLSIEBKCRLLTPHSAKSKRGYAKDVRSLSSKAVNH
HSWKDLLEDETETPIDTTIAKNSVFCVQPGKGRKALVFPDGLGRVCEKALVTR
VSVTLPOAVMGVSGFYQSPGHRVEFLVNAWKSXKMGKGAAYVDRCDSTVEDIRV
EESIYQCCDLAPEARQVIRSLTERLYVGGELTNSKNGCGRRLRCAAGVLTGAPNT
TCTYCSAAGCAALQDCTVLCGDLVLYCESGATGDEBSLSEFTEAMRVAAYSGPNT
DPPEBDETLITSCSNVSAHADSKRYVLYLRDPTPLAARAMEARLTPNVSMT
GNTIMYAPTLMARMLTMTFPSLILABOULEKMLDCOIVGATYSILPELDIPORTLH
GLTSFASHTSPGEINRVASCLRGVPLKRVHMRKASVRAKLSQGGRAATGCKYI
FNMAVRIKLTLPILPAASQDLDSGMPVAGISGGDIYHLSHARPMPFLMCLLLSVG
GIYILPVR"

```

ORIGIN	
Query Match	89.5%; Score 746.6; DB 14; Length 9379;
Best Local Similarity	95.1%; Pred. No. 1.1e-171;
Matches	792; Conservative 0; Mismatches 39; Indels 2; Gaps 2;
QY	1 AGAAAAACCAACGTAACCAACCTCCGCCCAAGAGCGTACGTTCCGGGGGGTGGT 60
DB	354 AGAAAAACCAACGTAACCAACCTCCGCCCAAGAGCGTACGTTCCGGGGGGTGGT 413
QY	61 CAGATCGTGGTAGATTACTCTGTGGCCGCGCAGGGGCCCCAGGTTGGGTGTGGCCGC 120



Db 414 CAGATCGTTGGGAGGTTTAACTGTTGGCCGCGACAGGGGCCCAAGGTGGGTGGCCGC 473

Qy 121 ACTAGAGAAGACTTCCGAGCGGTCGCAACCTCGTGGAGGGCGACACTATCCCAAGGCT 180

Db 474 ACTAGAGAAGACTTCTGAGCGGTCGCAACCTCGTGGAGGGCGACACTATCCCAAGGCT 533

Qy 181 CGCCAGACCCCGAGGGAGGGCCCTGAGCTACGCCCGGGATCCCTTGGCCCCCTCATGGAAC 240

Db 534 CGCCGCGCCGAGAGGGCAGGGCCCTGGCTCAGCCCGGGATCCCTTGGCCCCCTCATGGAAC 593

Qy 241 GAGGCGATGGGGTGGGGCAGAGATGGCTCTCTGTACACCCCGTGGCTCCCGGCTTATGGGGC 300

Db 594 GAGGCTCTGGGGTGGGGCAGAGATGGCTCTCTGTACACCCCGTGGCTCCCGGCTTATGGGGC 653

Qy 301 CCCATGACCCCGCGGAGTACGCTGCGGTATTTGGGAAAGTACGATGATACCTGACATGC 360

Db 654 CCCACAGACCCCGCGGAGTACGCTGCGGTATTTGGGAAAGTACGATGATACCTGACATGC 713

Qy 361 GGCCTTCGCGACTCTCATGGGGTACATTCGCTCGTGGCGCTCCCTTAGG-GGCGCTGC 419

Db 714 GGCCTTCGCGAC-CTCATGGGGTACATTCGCTCGTGGCGCGCCCTTAGAGAGGGCGTGC 772

Qy 420 CAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAGACGGCGTGAACCTATGCAACGGGAA 479

Db 773 CAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAGACGGCGTGAACCTATGCAACGGGAA 832

Qy 480 TTTACCCGGGTCCTCTTCTCATCTCCCTCTGGGCTTGGCTGCTGTTTACATTC 539

Db 833 TCTGCCCGGTGCTCTTCTCATCTCTCTCTGGGTTGGCTGCTGTTTACATTC 892

Qy 540 AGCTTCGCGTTATGAAGTGCAGACGATCCGGGATCTACATGTACAGAACGATTC 599

Db 893 AGCTTCGCGTTATGAAGTGCAGACGATCCGGGATCTACATGTACAGAACGATTC 952

Qy 600 CAATCAAGCATGCTGTACAGACAGCGGACATGATCATGCAACCCCGGGTGTGTGC 659

Db 953 CAATCAAGCATGCTGTGTAGAGCAGCGGACATGATCATGCAACCCCGGGTGTGTGC 1012

Qy 660 CTGTTCGCGGAGGGGTAATTCCTCCCGTGGGAGGCTCAGTCCCGAGCGTGGCGGC 719

Db 1013 CTGTTCGCGGAGGGGTAATTCCTCCCGTGGGAGGCTCAGTCCCGAGCGTGGCGGC 1072

Qy 720 CAABACGCGACGATCCCACTGCGACATACGACGCGCACTGATTTGCTGTTGGGCG 779

Db 1073 CAGGAATCGCACATCCCGGTTACGACATACGACGCGCACTGATTTGCTGTTGGGCG 1132

Qy 780 GGCCTCTTCTGCTCGCTATGATGATGCGGGGATCTCTGCGGATCTGTTTCC 832

Db 1133 AGCTGCTTCTGCTCGCTATGATGATGCGGGGATCTCTGCGGATCTGTTTCC 1185

RESULT 13

AF165051 9379 bp RNA linear VRL 04-NOV-1999

LOCUS Hepatitis C virus strain MD4-1 complete genome.

DEFINITION AF165051

ACCESSION AF165051.1 GI:5918940

VERSION

KEYWORDS

SOURCE

ORGANISM

Hepatitis C virus

Hepatitis C virus

Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE

1 (bases 1 to 9379)

AUTHORS Nagayama,K., Kurotaki,M., Enomoto,N., Maekawa,S.Y., Miyasaka,Y., Tazawa,J., Izumi,N., Marumo,F. and Sato,C.

TITLE Time-related changes in full-length hepatitis C virus sequences and hepatitis activity

JOURNAL Virology 263 (1), 244-253 (1999)

MEDLINE 20013325

PUBMED 10544098

REFERENCE 2 (bases 1 to 9379)

AUTHORS Nagayama,K., Kurotaki,M., Enomoto,N., Maekawa,S., Miyasaka,Y.,

TITLE  
JOURNAL  
FEATURES  
SOURCE

Sakamoto,N., Fukuma,T., Tazawa,J., Izumi,N., Marumo,F. and Sato,C.  
Direct Submission  
Submitted (06-Nov-1999) Second Department of Internal Medicine,  
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-Ku,  
Tokyo 113-8519, Japan

CDS

1..9379  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/strain="MD4-1"  
/db\_xref="taxon:1103"  
/note="from a patient with persistently normal serum ALT values; patient 4, point 1  
genotype: 1b"  
330..9362  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="A056186.1"  
/db\_xref="GI:5918941"  
/translation="MSTNPKQRKTKNTNRPODVKPPGGQIVGGVYLIPRGRPL  
GVARTRKTSERSQPRGRROPJPKAREBGRWMAOPGYPMPLYGNEGGMWMLSPRG  
SRPMGPTDPRRRSRNLGKVIDTLTCGFADIMGYIPLVGAFLGGAALAGVLELD  
GVNVAATNLPGCSRSIFLAILSLCTIPASAEVYRNSGIYHTVNDGNSISIVAAD  
VIMTPGCVPEVRSNSRCKVALTPTLAANASISPTTTPRHVDILVGTARPCSAAY  
VGDLCGSPFLISQ.LFTSPRHEIVQDNCSTYGHVGRHMDMMNMSPTLALV  
SOLIRIQAIVDVMAAGWGLAAGLAAVYAGNNAKAVLIVLLEAGVAGNTYTSGLV  
GROTSGLTSLFVSAGSOKIQLINTNGSMHINRTANCNDSIQTFPLALETYHFRFMS  
GCPRMASCRPIDFAQMGPIITYAEPSPORRYCHMAKPCGIIYASVSCPVYC  
PTSPVAVGTTDRSGVPTVYMGENTVLLNTRPPQGNMGCTYMAATFTGTCGG  
PPCHIGVGNNTLICPTDCKPKHBEATYTKGSSPMVTPRITVYPRILMYPCPTNV  
TIFVRMYVGVGEHRLNAACMWTGECRDLDRSELSPILSITTEKQILPCPTTL  
PALSTG.IHLHQNVQVLYGLISAVSFAIKMEVYL.LFLIADARVCA.LMMMLL  
IAQAEALENLVMDENASVAAHGLISFLVFCAAMYIKGLVGAAYAIYGVPLIL  
LIL.LPSRAVYMDENASVCGGVEFGALITLSPYKFLARLMMMLLOYEITPAEAM  
LOWVAPLNRGGRDAIILACAVHPELIPITKILISLIGGALASIVRVPFAR  
AOGIIRACMLYRAKAGCHYACMVYRLAALTLGYVYDHLPLRMAHAGLDAVAYE  
PVSVIDETKVTWGAADTACGDITISGLPVARGKREILGPAOSIBOGRRLAPIT  
AYSQOTGILCCLITSLTGRDNKVEBQVVSATQSFATCVNGACVYHFAHAGSK  
TLAPKGPITQMYTNVQDILVGMQAPARSLPTCTGSSLDVYVTRHADVIPYRGR  
DTRGSLSPRSVSLKSSGGFLPCSHAVGIFRAAVCTRGVAKA.VDFIPEVAMEET  
MRSPVPTDSSPPAVPOTPOVYAHAAHPGSGSKRPVAAVAGCYKVLNPSVAATL  
GFGAYMKASHTEENIRGTVATTITTGAPITTYSTYTKPLADGCGSGAGV.DIICCECHS  
TDTSTIILIGITVLDQAEAGARLVLAATPAGSVTPVPHNIEVALSNSEIIPYKG  
AIPLEVIKGRNHLIFCHSKKCDLAKLSLGLNAAYYXGLDVSVIPSTGDDVVVA  
TDALMTGFTGDPDSVIDNCTVTVTFVPSLPLFTIETTVPODAVRSQRGRGTCG  
RSGIYFVTPGERSGMPDSVLCECDAGCAMELPAETSVSLRALVTPGLPVCQ  
DHLSPWSVPTGLYHDAHPT.SOTKQAGDNPPYVAYOATYCARAOPAPSPMDWKC  
LIRKPLTHGPTPLYLHGAAYQNEVTLTHPTIKTIMCMSADLBYVSTWLVGVYLA  
ALAAVCLITTSVIVGRIILSGRAVILPDRVLIREPEMECASHLPYIQGQJLAB  
QPKQKALGLQATATKQEAAPVVESSKWRLEAEAWAKMMNPIG.IOYLAQSLTLPN  
PAISLMAFTASITSPITTOSTLLENILGAVNAQLAPPSAASAFVAGIAGA.VGSI  
GLGVYLVNLIAGYAGVAGALVAFKINSVESTEDVNLPAISLPAALVGVCA  
ILRRVHVGEGEAGVOMNRILTA.PASRGHNVSPHYVPESDAARVTOILSLITQILK  
RLHQMINEDCSTPCSGSWLRDVMWITCVLSDPRTWLOSLLPLRBLPI.SSCQRYK  
GWMRGDEIMQTTCCGAGQITGHRVNGSRVITGPPKSCVWGTGTPFINATYTGCPCLR  
APNYSRALMRVAAEYVEVTVGDEPHVNTKICPPCTVAPAEPEFTEVDEGRPLR  
YARACKPLIREVTFQVGLNOYLIVGSQLPCPEPDDVAVLTLM.LDPSHIIETEARL  
LAGBPPLSIASSASQISAPSLKATCTTGNDPDDLIEANLMEOMSGANTITRESSE  
KVYLIDSTPILRAEDEREVSVPAILKRSKRPMAWVMPARPDYNPIESAWDPOY  
VPPVHGGCPLEPTAPPIPPRRKRVTLTSTYSALAEIATTFGSSGSGASDSTG  
ATAPPDQASDDGKSDVESYSMPLEBEGDPLDLSAGSIVSEESSESVCCSMS  
YTMWGLITTPCAAEESKPLINSLSNLRIHNMVAYATTSRSAROKRVTPRGLQVLD  
DHYRDLVKEMKASATYKAKLSIEEAKCLTPPSAKSKFGYALRQKLVNLTGKLVNLI  
RSWKDLLEDETEPIIDTTVMAKNEVPIOPKESKRPVAVPDIAGRVCEKMAIYD  
VSTPLPQAVMSSSTGFQYSPKQRYEFLVMAKSKCKPMGRSVDTRCEPSTYENDIRV  
EESIQCCDLA.PEARQAKSLTERLYIGGLTNSKGQCGYRRCASVLTSCGNTL  
TCYKASAAACRAAKIQCCTMUNGDVILVIESGOTDANLRFETAMRYASAPG  
DPPQAYDLELTSCSSNVSAHADSGKRVYILTRDPTTPARAAMETARHTPVWSML  
GNITMYAPTLMAWVLMTHFRSILLAOEOLKALDLCQIYGAVYSIEPDLQPIERLH  
GLSAFSLHSYSPGEINRYDAACLRKLGVPLELVWHRASVYAKLSISQGRALTCCKYL  
ENMAVTRTKLTPIPASASQDLDSWFAVIGSGDIYHLSLRAPRMFWLCLILSVGV  
GIYLLPNR"

ORIGIN



Query Match 89.3%; Score 745; DB 14; Length 9379;  
 Best Local Similarity 95.0%; Pared. No. 2, 8e-171;  
 Matches 791; Conservative 0; Mismatches 40; Indels 2; Gaps 2;

```

QY 1 AGAAAAACCAACGTAACCAACCACTCCGCCACAGACGTCAGGTTCCGGGCGGTGGT 60
DB 354 AGAAAAACCAACGTAACCAACCAACCCGCCACAGACGTCAGGTTCCGGGCGGTGGT 413
QY 61 CAGATCGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
DB 414 CAGATCGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 473
QY 121 ACTAGGAAGACTTCCGAGCGGTGCGAACCTGCGAAGGCGCAACCTATCCCAAGGCT 180
DB 474 ACTAGGAAGACTTCCGAGCGGTGCGAACCTGCGAAGGCGCAACCTATCCCAAGGCT 533
QY 181 GCGCAGCGCCGAGGCGAGGCGTGGGCTCAGCGCGGCTACCTTGGCCCTTATGGCAAC 240
DB 534 GCGCGGCGCGAGGCGAGGCGTGGGCTCAGCGCGGCTACCTTGGCCCTTATGGCAAC 593
QY 241 GAGGCGATGGGCGTGGGCGAGATGGCTTCTGTCAACCCCGTGGCTCCGGGCTATGGGCG 300
DB 594 GAGGCGATGGGCGTGGGCGAGATGGCTTCTGTCAACCCCGTGGCTCCGGGCTATGGGCG 653
QY 301 CCCACTGACCCCGGCGGTAGGTGCGGTAATTTGGGTAAGTATCATGATACCTTACATGCG 360
DB 654 CCCACTGACCCCGGCGGTAGGTGCGGTAATTTGGGTAAGTATCATGATACCTTACATGCG 713
QY 361 GGGTTCGCGGCACTCTCATGGGTAATTCGCTCGTGGCGCTCCCTTA-GGGGCGCTGG 419
DB 714 GGGTTCGCGGCACTCTCATGGGTAATTCGCTCGTGGCGCTCCCTTAAGGGGCGCTGG 772
QY 420 CAGGCGCTTGGCGGCGATGGGCTCCGGGTTTGGAGGAGCGGCGTGAATATGCAAGGAA 479
DB 773 CAGGCGCTTGGCGGCGATGGGCTCCGGGTTTGGAGGAGCGGCGTGAATATGCAAGGAA 832
QY 480 TTTTACCGGCTGCTCTTCTCTATCTTCTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 539
DB 833 TTTTACCGGCTGCTCTTCTCTATCTTCTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 892
QY 540 AGCTTCGCGTATGAAAGTGCGCAACGTCGCGGATCTACATGTCAGCAACGATTTGCTC 599
DB 893 AGCTTCGCGTATGAAAGTGCGCAACGTCGCGGATCTACATGTCAGCAACGATTTGCTC 952
QY 600 CAATTAACGATCTGTATGAGAGCAGCGGACATGATGACACACCCCGGGTGTGCTGCC 659
DB 953 CAATTAACGATCTGTATGAGAGCAGCGGACATGATGACACACCCCGGGTGTGCTGCC 1012
QY 660 CTGTGCGCGGAGGGAATTTCTCCGCGCTGCGGATGAGGCTCATCTCCACGCTCGCGGC 719
DB 1013 CTGTGCGCGGAGGGAATTTCTCCGCGCTGCGGATGAGGCTCATCTCCACGCTCGCGGC 1072
QY 720 CAGAGCGCGAGCATCCCACTGCGCAATACGACGCCAGTCGATTTGCTCGTTGGGCG 779
DB 1073 CAGAGCGCGAGCATCCCACTGCGCAATACGACGCCAGTCGATTTGCTCGTTGGGCG 1132
QY 780 GGGTGCCTTCTGCTCGCTATGATGCTGGGGGATCTCTGCGGATCTGTTTCC 832
DB 1133 GGGTGCCTTCTGCTCGCTATGATGCTGGGGGATCTCTGCGGATCTGTTTCC 1185

```

RESULT 14  
 AF165052  
 LOCUS AF165052 9379 bp RNA linear VRL 04-NOV-1999  
 DEFINITION Hepatitis C virus strain MDA-2 complete genome.  
 ACCESSION AF165052  
 VERSION AF165052.1 GI:5918942  
 KEYWORDS  
 SOURCE Hepatitis C virus  
 ORGANISM  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.

## REFERENCE

1 (bases 1 to 9379)  
 Nagayama, K., Kuroski, M., Enomoto, N., Maekawa, S. Y., Miyasaka, Y.,  
 Tazawa, J., Izumi, N., Marumo, F. and Sato, C.  
 Time-related changes in full-length hepatitis C virus sequences and  
 hepatitis activity  
 Virology 263 (1), 244-253 (1999)  
 2 (bases 1 to 9379)  
 Nagayama, K., Kuroski, M., Enomoto, N., Maekawa, S., Miyasaka, Y.,  
 Sakamoto, N., Fukuma, T., Tazawa, J., Izumi, N., Marumo, F. and Sato, C.  
 Direct Submission  
 Submitted (06-JUL-1999) Second Department of Internal Medicine,  
 Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,  
 Tokyo 113-8519, Japan

## FEATURES

## source

## CDS

```

/codon_start=1
/product="polyprotein"
/protein_id="ABD56187.1"
/db_xref="GI:5918943"
/translation="MSTNPKPQKTRNTNRAPDQVKEPGGQIVGVVLLPRGRRL
GVARTKTSERQPRGRQPIPKARPERKTAQVGPMPILGNGMGAGMILSRG
SRPSMGPTDPRRRSRNLKVIDLTGDFADNGYIPLVAGPLGGAARALAHGVLEAD
GVNATNGLPGCSPSIFILLALSCITPASAYEVNNGVHYTNDGSSSIYVEAD
VIMHPGVCPEVRESNSRCVATLPLAARNAVPTTIRRHVDLVGTAFCSAMY
VGDLCGVFLISLFTPSRRHETVDNCSTIYPHVSQHRAMMMNMMSPITALAV
SOLRLIPOAVDMVAGAWGYLAGLAVYVSMANMAKVLVMLLFGVDEITHTGAV
SHRTFRASLFSKASOKIOLINTGSMWINTALNCNDLOTGLTGLFVYRFPAS
GCRPVAACRPIDKFAQGMGPIITYAHSPPDRPCMHAPRCPGVPSSEVCGPYC
PTSPSVGACGTRDPSGVPTTYGNGENTDILLNTRPQNGMGCTYMANATGTRKCG
PPCHIGGVGNNTLICTPCFRKPREATTYKCSGMLTPRCIVDPIRYMHYPCYANP
TIFKRVAVVGVHRLNACNWTRECDLERDSELSPLLSTTEWQIILICDECHS
PALSTGLIHLHONIADVOLYXIGISAVYFAIKMEVYLIFLLADARVACILMMFL
IAQBAALENLVILNAASVARRAHGILSLFPCAMVYIKGLVPGAAVYIYGVMPLL
ILLALPRAVYANDREMAASCGGAVGVALLTLSPYKVELRLMMWLOYPITRAEH
LOWMPPLANRGRARAILIACAVBELIFDTIKLLSITGLMLOASIVAVPYVR
AQLIRACMLVRKAGAGHTAQVLAALATITYYDHLPIQDNHAGELDLAVTE
PVFSDMETKXVITWGAADTAAGDIISGLPVSARKREILIGADNVEGGKMLAIVT
AYSQOTRGLGCIITSLGRDNQVGEVQVSTATOSFLATCVGACATVFRHAGSK
TLAEPKGPITOMYTVNDODLVGMOAPPGARSILPTCGSGSDLYLVTRHADVIPIVRG
DTRGSLSPRPVSYLKSGSGPILPDSGAVGI PRAAVCTRGVAKAVDIEIPIVAMETT
MNSPVPTDNSPSPANVPTQVNHIAFPSSGKSTVPYAAAGCYKIVLVNVSVAATL
GFGAYMSKAGHTEPNIRITGVRTITTGAPITSTYTGFLADGGCGGADYIILICDECHS
TDSITLIGITVLDQAEAGARLVVLAATPFGASVTPHPNIEVALSTWGIIPFYGK
AIPILVIKGRHLIFCHSKKCDLELAKSALGLAAVAYRGLDVSVIPSGDVVVA
TDALMTGTFDESDVIDCNTCVTQVDSLDPTPIETTTETVODAVSREORGTRG
RGCIYFVTPGERSGMDSSVLCQVYAGCAWELTAFETSRLRAYINTGELPVCQ
DHLPEWESYFTGLTHIDAFISQTOAGDNFPIYVAQATVAKAQAPEPSDQMKC
LIRLPTLHGPPLIRUGAVONEVTLPIKFLIMACASALVEVTSIYVIVGGLA
LAAVCLTGGSVVIVGRIILSGRPVAVIPRELYREPDMECASHLPIREGMGLAE
QFKQALGILQIATQAEAAAPVSESKRALFAKAKMMNTISGIVYLAGISTLPGN
PALASIMAFATASITPLTOSTLIFNLIGAWAOLAPASASAFAGIADAAGVSI
GIGVLYNDIAGVAGVAGVAVAPKIMSGEVSPTEDIVNLLPAILISPGALVGVCA
ILRRHYREBQAVQMMRLIAPASGNHVSPTHYIPESDAARVTOILSLITLQILK
RLHQMINEDCSTPSCGSLRDVDMWICVLSDFKMLQSLKPLRGLPGLSCQRYK
GWRVGDIMQITPCGAQITGVHKNKSNRIVPCKTCSNMWHTPIFNATVTPCPS
APNYSRALMRVAAEEVETVRGDEHYVGMNDNIIKCCQVPAEFPEFVGVDLHR
YAPACKPLIREEVTPOVGNOLYVGSOLPCEBEPVAVVITSLTSPHSTATFARRL
AAGSPSLASSASASOLASPSLAKCTTTHDSDDADLIANLILMRMGENTRSEN
KVYIIDSDEPLAEEDERESPVAITLKSRRFPNAMPYMAPPDNPLPLESKAPDY
VPPVHAGCLPPTRAPLIPPPRRKRTVLTSTVSALAEIKTFSGSSGSAVSGS
ATAAPDOASDDDGSDVESYSMPLREBEPDPLSDSWSVTEBESGSDVCSMS
YWTGALITPCAABESKLPINLSNLSLRHNMVYATTSRSLQKQKVTPERLQVLD
DHYRDLKEMKAKASTVAKKLISIEACKLTPPHSAKSKFGVAGADVRLTGKAVNHI

```



RAGIYRVTYPERSPGMPDSVSLCECYDAGCAMYEITPAETSVRLRAYLNTBGLPVMQ  
 DLHEFMSVFETLTHIDAHFLSOTKOAGNFPYLYAVOATVCARAOAPPPSDOMKIC  
 LIRKPTLHGPTLPIYLGAVONNEVILITPTIKYIMACASADLEVTSTWVGVGLA  
 ALAACTTSGSVIVYGRITILSKPAIIPDRVLVREPMECSAIPYIEGOMLAE  
 QPKQALGLOLATKOAEAAAPVESKMOALTTPAKMMNFISGIIQIAGSTIPGN  
 PAIASLAFASITSPLTTOHTLFLNIIIGGVAAQLAPPRASAFVAGIAGAAGSI  
 GIGKVLVDILAGYGAGVAGALVAFVMSGDMSTEDLVLLPAIISPGALVVGVCIA  
 ILRHVGEAGVOMNRILAFASRGNVSPHYVESPAARVTOISLITLOLK  
 RLHOMINDCSTPGSMILRDYMDICTYLSDFKTMLOSRLVPLRGDPVPECSOGY  
 GWMRDGIMQTTCCGCAQITGHVNGSRKIVPCTSNMHTFPINATTPCPSPP  
 APTSRALMRVAEEIVEYTRVGDHYVTGMTDNUKCCQCPABEFKELGCVLHR  
 YAPSKPLRDEVTQVGLNOYVGSOLCEPEPVALVTSMLTDPESHITAEARSL  
 ARGPSPSLASSASOLASPLKATCTTHDSPADLIEANILMRGEMGNITRVSEN  
 KVLIDSEPLAEEDEREVSLPAELIKSRKFPAMPIMAPDYNPLIEEMQDY  
 VPPVHGGLPPTKAPPIPPRRKRTVLTSTVSALAELATKTFGSSSSAASGT  
 ATAEPDOPSSDDAGSDVESYSMPLEBEPDPLSDGSMSTVSEBAGEDVCCMS  
 YTWGLITPCAAESSKLPIINALSNSLLHNMVYATTRSASOROKVTFRLOVLD  
 DHYRDLKEMKAKASTVKARLISVEACKLTPHARSFKFGAGADVRLSRAINLI  
 RSVMDLOBDETPTIDITIMAKNEVCVQPEKGRKPAALIVPDLGVYVCKMALYD  
 VYSTLPQAVMGASVGYOYSGQREVELVNAKSKCPMGFAVDTCFDSVTENDIRV  
 ESISTOCDDIAPEAROALRSILTERLYIGPLNSKGNONGVRCASGVLITSCNTL  
 TCYLKATACRAKIQDCTMLVCGDPLVITCESACTQEBPABALRPFTBAMTRYSPG  
 DPEPEYDELITSCSNVSVADHDSGRKVVYLTDPPTPLARAAEETARHPVNSWL  
 GNIIMVAPTLMARMLTMTHFSLILAOBOLKALDQOIGACYSIEPLDPLQIIRLH  
 GLSARSLHSYSGEINRVASCRLGVPLRWRRHARSRAKLLSOGGRAALCGKYL  
 FWMVAVRTKLTPIPASQLDLSGMFVAGYSGDLYHVSRRARPRFMWCLLSVGV  
 GYVLNPR"

mat\_peptide 330..902  
 /product="core protein"  
 mat\_peptide 903..1478  
 /product="E1"  
 mat\_peptide 1479..2756  
 /product="E2"  
 mat\_peptide 2757..3407  
 /product="NS2"  
 mat\_peptide 3408..5300  
 /product="NS3"  
 mat\_peptide 5301..5462  
 /product="NS4A"  
 mat\_peptide 5463..6245  
 /product="NS4B"  
 mat\_peptide 6246..7586  
 /product="NS5A"  
 mat\_peptide 7587..9359  
 /product="NS5B"  
 misc\_feature 1479..1559  
 /note="hypervariable region (HVR)"  
 misc\_feature 6954..7073  
 /note="interferon sensitivity determining region (ISDR)"

## ORIGIN

Query Match 89.3%; Score 745; DB 14; Length 9410;  
 Best Local Similarity 95.0%; Pred. No. 2.8e-171;  
 Matches 791; Conservative 0; Mismatches 40; Indels 2; Gaps 2;

QY 1 AAAAAAAAAAAGCTTACCACTCCGCCCAAGAGCTGCTCCCGGGCGGTGGT 60  
 DB 354 AAAAAAAAAAAGCTTACCACTCCGCCCAAGAGCTGCTCCCGGGCGGTGGT 413  
 QY 61 CAGATCGTGTGAGATTACCTGTTCCGCGCAGAGGGGCCCAAGTTGGGTGTCCGCG 120  
 DB 414 CAGATCGTGTGAGATTACCTGTTCCGCGCAGAGGGGCCCAAGTTGGGTGTCCGCG 473  
 QY 121 ACTAGGAAGACTTCCGAGCGGTGCAACTCTGTGGAAGGAGCAACTATCCCAAGGCT 180  
 DB 474 ACTAGGAAGACTTCCGAGCGGTGCAACTCTGTGGAAGGAGCAACTATCCCAAGGCT 533  
 QY 181 CGCCAGCCCGAGAGGCGAGGCGTGGCTACGCCCGGGTACCTTGCGCCCTCTATGGCAAC 240  
 DB 534 CGCCAGCCCGAGAGGCGAGGCGTGGCTACGCCCGGGTACCTTGCGCCCTCTATGGCAAC 293  
 QY 241 GAGGCGATGGGCTGAGAGGCTCTCTGTCACCCCGTGGCTCCGCGCTACTGTTGGGCG 300  
 DB 741 GAGGCGATGGGCTGAGAGGCTCTCTGTCACCCCGTGGCTCCGCGCTACTGTTGGGCG 803

DB 594 GAGGCGATGGGCTGAGAGGCTCTCTGTCACCCCGGCGCTCTCGGCTTACTGTTGGGCG 653  
 QY 301 CCCACTGACCCCGGCGTAGGTGCGGTAAATTTGGGTAAAGTCATGATACCTTCATGCG 360  
 DB 654 CCCACTGACCCCGGCGTAGGTGCGGTAAATTTGGGTAAAGTCATGATACCTTCATGCG 713  
 QY 361 GCGTTCGCGCACTTCATGAGGGGTACATTCCGCTGCTGCGGCGCTCCCTTA-GGGCGGCTGC 419  
 DB 714 GCGTTCGCGCACTTCATGAGGGGTACATTCCGCTGCTGCGGCGCTCCCTTA-GGGCGGCTGC 472  
 QY 420 CAGGCGCTGCGGCGATGCGCTCCGCGGTTCGAGAGACGCGCGTAATGACAAACAGGGA 479  
 DB 773 CAGGCGCTGCGGCGATGCGCTCCGCGGTTCGAGAGACGCGCGTAATGACAAACAGGGA 832  
 QY 480 TTACCGGCTGCTCTTCTCTATCTTCTCTTGGCTTGTGCTGCTGTTGACCATTC 539  
 DB 833 TCTGCGCGGTCTCTTCTCTATCTTCTCTTGGCTTGTGCTGCTGTTGACCATTC 892  
 QY 540 ACCTTCGCTTATGAGTGCAGCAAGCTGTCGCGGATCTACATGTCACGACGATTGCTC 599  
 DB 893 ACCTTCGCTTATGAGTGCAGCAAGCTGTCGCGGATCTACATGTCACGACGATTGCTC 952  
 QY 600 CAATCAGCATCGTGTACGAGACAGCGGACATGATCATGCAACCCCCGGGTGTGTCG 659  
 DB 953 TAACCGACATGTTGTATGAGGCGAGCGGACATGATCATGCAACCCCCGGGTGTGTCG 1012  
 QY 660 CTGTGCGGAGGGAATTCTCCCGGCTGCTGGGAGGCTCACTCCAGCGCTCGCGCG 719  
 DB 1013 CTGTGCGGAGGGAATTCTCCCGGCTGCTGGGAGGCTCACTCCAGCGCTCGCGCG 1072  
 QY 720 CAGAGCGCAGCATCCCACTGCGACATACAGAGCGCAAGTTCGTTGGGGC 779  
 DB 1073 CAGAGCGCAGCATCCCACTGCGACATACAGAGCGCAAGTTCGTTGGGGC 1132  
 QY 780 GCGTCTTCTGCTGCTGCTATGATGAGTGGGGGATCTTCGCGATCTGTTTTC 832  
 DB 1133 GCGTCTTCTGCTGCTGCTATGATGAGTGGGGGATCTTCGCGATCTGTTTTC 1185

Search completed: February 21, 2005, 03:06:16  
 Job time : 4119 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2005, 21:58:35 ; Search time 588 Seconds  
(without alignments)  
8396.368 Million cell updates/sec

Title: US-09-664-363-5

Perfect score: 834  
Sequence: 1 AGAAACCAACGTAACAC.....TCTGGGATCTGTTTCCCG 834

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1808:\*  
2: geneseqn19908:\*  
3: geneseqn20008:\*  
4: geneseqn20018:\*  
5: geneseqn20018s:\*  
6: geneseqn20028s:\*  
7: geneseqn20028s:\*  
8: geneseqn20038s:\*  
9: geneseqn20038s:\*  
10: geneseqn20038s:\*  
11: geneseqn20038s:\*  
12: geneseqn20048s:\*  
13: geneseqn20048s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	834	100.0	834	2	AAQ12238
2	804.8	96.5	2790	2	AAQ46191
3	804.8	96.5	3372	2	AAQ46202
4	801.6	96.1	831	2	AAQ46192
5	799.4	95.9	2116	2	AAQ12242
6	738.6	88.6	1734	2	AAQ40428
7	738.6	88.6	1734	2	AAQ40429
8	737	88.4	1734	2	AAQ40430
9	737	88.4	1734	2	AAQ40438
10	737	88.4	2433	2	AAQ12974
11	737	88.4	2433	10	ADDS5557
12	737	88.4	2433	12	ADP71139
13	737	88.4	9587	13	ADR82189
14	733.8	88.0	2187	2	ABAQ3491
15	733.8	88.0	2560	2	AAQ43889
16	733.8	88.0	3360	2	AAQ43877
17	733.8	88.0	3461	2	AAQ43867
18	733.8	88.0	3461	2	AAQ43868
19	733.8	88.0	9413	2	AAQ43869
20	733.8	88.0	9413	2	AAQ43870

21	733.8	88.0	9413	6	AAQ25517	AAQ25517 Hepatitis
22	733.8	88.0	9413	8	AAQ53723	AAQ53723 Hepatitis
23	733.8	88.0	9413	8	AAQ49655	AAQ49655 Hepatitis
24	733.8	88.0	9413	10	ADP88596	ADP88596 Hepatitis
25	733.8	88.0	9611	13	ADQ34713	ADQ34713 Hepatitis
26	733.8	88.0	1880	2	AAQ24467	AAQ24467 NANB hepa
27	732.2	87.8	1880	2	AAQ63753	AAQ63753 NANBH ge
28	732.2	87.8	2540	2	AAQ64069	AAQ64069 Non-A, no
29	732.2	87.8	3401	2	AAQ30387	AAQ30387 5'UTR/COR
30	730.6	87.6	1270	2	AAQ60658	AAQ60658 Fragment
31	730.6	87.6	1562	2	AAQ60672	AAQ60672 Fragment
32	730.6	87.6	1734	2	AAQ40437	AAQ40437 Hepatitis
33	730.6	87.6	1953	8	AAQ55222	AAQ55222 Plasmid P
34	730.6	87.6	2540	2	AAQ29628	AAQ29628 Hepatitis
35	730.6	87.6	2829	2	AAQ60673	AAQ60673 Fragment
36	730.6	87.6	9413	2	AAQ80498	AAQ80498 DNA encod
37	730.6	87.6	9595	2	AAQ24843	AAQ24843 Infectiou
38	730.6	87.6	9595	4	AAQ23492	AAQ23492 Infectiou
39	730.6	87.6	9595	4	AAQ86939	AAQ86939 Nucleotid
40	730.6	87.6	9595	12	ADQ36222	ADQ36222 Hepatitis
41	730.6	87.6	9599	12	ADQ79396	ADQ79396 Hepatitis
42	730.6	87.6	9599	2	AAQ24833	AAQ24833 Infectiou
43	730.6	87.6	9605	6	ABK91431	ABK91431 Hepatitis
44	730.6	87.6	9605	6	ABK91424	ABK91424 Hepatitis
45	730.6	87.6	9605	6	ABK91429	ABK91429 Hepatitis

## ALIGNMENTS

RESULT 1	AAQ12238	standard; DNA; 834 BP.
ID	AAQ12238	
XX	AAQ12238;	
AC	XX	
XX	XX	
DT	25-MAR-2003 (revised)	
DT	06-SEP-1991 (first entry)	
XX	XX	
DE	Clone Brl1 encoding PT-NANBH virus antigenic portion.	
XX	XX	
KW	post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.	
XX	XX	
OS	Non-A.	
OS	non-B hepatitis virus.	
PN	GB2239245-A.	
XX	XX	
PD	26-JUN-1991.	
XX	XX	
PF	17-DEC-1990; 90GB-00027250.	
XX	XX	
PR	18-DEC-1989; 89GB-00028562.	
PR	27-FEB-1990; 90GB-00004414.	
PR	03-MAR-1990; 90GB-00004814.	
XX	XX	
PA	(WELL ) WELLCOME FOUND LTD.	
PA	(HIGH/) HIGHFIELD P E.	
PI	Highfield PE, Rodgers BC, Tedder RS, Barbara JW;	
DR	WPI, 1991-187564/26.	
DR	P-PSDB; AAR12596.	
XX	XX	
PT	Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA	
PT	and antibodies used in diagnostic assays and in vaccines.	
XX	XX	
PS	Claim 10; Page 56-58; 108bp; English.	
XX	XX	
CC	This sequence is a structural region of the PT-NANBH viral genome	
CC	encoding an antigenic polypeptide. It was isolated from serum A CDNA	
CC	library was prepared in lambda gtl1 from the serum of infected patients	
CC	and screened with antibodies from the serum of humans with a high risk	

CC for PT-NANBH but which did not react with viral antigens DX113, BHC-5 and  
 CC BHC-7. Clone BR11 was identified which did not cross-hybridize with  
 CC probes made from JG2 and JG3 (see AAQ12236 and AAQ12237). It was  
 CC sequenced and found to have the sequence shown, which includes the EcoRI  
 CC linkers added during cloning. See also AAQ12239-Q12242. (Updated on 25-  
 CC MAR-2003 to correct PA field.)

XX Sequence 834 BP; 139 A; 265 C; 251 G; 179 T; 0 U; 0 Other;

Query Match 100.0%; Score 834; DB 2; Length 834;

Best Local Similarity 100.0%; Pred. No. 3.9e-216; Mismatches 0; Indels 0; Gaps 0;

Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAAAACCAAACTGTAACCAACCACTCCGCCCAGAGAGCTGAGGTTCCCGGGCGGTGCT 60  
 DB 1 AGAAAAACCAAACTGTAACCAACCACTCCGCCCAGAGAGCTGAGGTTCCCGGGCGGTGCT 60  
 QY 61 CAGATCGTTGGTGGAGTTTACCTGTTGCCGCGCAAGGGGCCCAAGTTGGGTGGCGCG 120  
 DB 61 CAGATCGTTGGTGGAGTTTACCTGTTGCCGCGCAAGGGGCCCAAGTTGGGTGGCGCG 120  
 QY 121 ACTAGAAAGACTTCGAGAGCGTGCACCTCGTGGAAAGGCGCAACCTATCCCAAGGCT 180  
 DB 121 ACTAGAAAGACTTCGAGAGCGTGCACCTCGTGGAAAGGCGCAACCTATCCCAAGGCT 180  
 QY 181 CGCCAGCCCGAGAGGCGAGGCGCTGGGCTCAGCCCGGGGTACCTTTGGCCCTCTATGCAAC 240  
 DB 181 CGCCAGCCCGAGAGGCGAGGCGCTGGGCTCAGCCCGGGGTACCTTTGGCCCTCTATGCAAC 240  
 QY 241 GAGGGCATTGGGTGGGCGAGATGCTCTGTACACCCCGGTGCTCCCGGCTATGTTGGGCG 300  
 DB 241 GAGGGCATTGGGTGGGCGAGATGCTCTGTACACCCCGGTGCTCCCGGCTATGTTGGGCG 300  
 QY 301 CCCACTGACCCCGCGGCGTGAAGTGGGTAATTGGGTAAGTGAATGATGATACCTCAATGC 360  
 DB 301 CCCACTGACCCCGCGGCGTGAAGTGGGTAATTGGGTAAGTGAATGATGATACCTCAATGC 360  
 QY 361 GGGCTTCGCGACTCTCATGAGGTTATCCGCTCGTGGCGCTCCCTTAAGGGCGCTGCC 420  
 DB 361 GGGCTTCGCGACTCTCATGAGGTTATCCGCTCGTGGCGCTCCCTTAAGGGCGCTGCC 420  
 QY 421 AGGGCCCTGGCGCATGCGGCTCCGGGTCTGGAGAGCGGGTGAACCTATGCAACAGGGAAT 480  
 DB 421 AGGGCCCTGGCGCATGCGGCTCCGGGTCTGGAGAGCGGGTGAACCTATGCAACAGGGAAT 480  
 QY 481 TTACCCGGGTGCTCTTCTCTATCTTCTCTGCTTGGCTTGGCTGCTGTTGACATTCGA 540  
 DB 481 TTACCCGGGTGCTCTTCTCTATCTTCTCTGCTTGGCTTGGCTGCTGTTGACATTCGA 540  
 QY 541 GCTTCGCGTTATGAGTGGCAACGTTGTCGGGATCTACATGTACGAAACGATTGCTCC 600  
 DB 541 GCTTCGCGTTATGAGTGGCAACGTTGTCGGGATCTACATGTACGAAACGATTGCTCC 600  
 QY 601 AACTCAGCATGTTGTAAGAGACAGCGGACATGATCATGACACCCCGGGGTGTGTGCC 660  
 DB 601 AACTCAGCATGTTGTAAGAGACAGCGGACATGATCATGACACCCCGGGGTGTGTGCC 660  
 QY 661 TGTGTCGGGAGGTAATTCCCTCCGCTGGGTGGTGAAGGCTCACTCCACGTCGCGGCC 720  
 DB 661 TGTGTCGGGAGGTAATTCCCTCCGCTGGGTGGTGAAGGCTCACTCCACGTCGCGGCC 720  
 QY 721 AAGGACGCGAGCATCCCACTGCGACAAATACGACGCCAAGTGGCTCTGTTGGGCG 780  
 DB 721 AAGGACGCGAGCATCCCACTGCGACAAATACGACGCCAAGTGGCTCTGTTGGGCG 780  
 QY 781 GCTGCTTCTGTCGCGCTATGTAAGTGGGGAATCTCTGCGGATCTGTTTCCCG 834  
 DB 781 GCTGCTTCTGTCGCGCTATGTAAGTGGGGAATCTCTGCGGATCTGTTTCCCG 834

RESULT 2  
 AAQ46191  
 ID AAQ46191 standard; DNA; 2790 BP.

XX AAQ46191;  
 AC 27-AUG-2003 (revised)  
 XX 25-MAR-2003 (revised)  
 DT 24-FEB-1994 (first entry)  
 DE PT-NANBH virus BHC-11 fusion protein.  
 XX  
 XX Parenterally transmitted non A non B hepatitis; PT-NANBH;  
 KM hepatitis C virus; HCV; NS5; E1; linker; detection; diagnosis; antigen;  
 KM vaccine; BHC-11; replicase; core protein;  
 KM Autographa californica nuclear polyhedrosis virus; AcNPV; polyhedrin; ss.  
 Hepatitis virus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2790  
 FT /tag= a  
 FT /product= "fusion\_protein"  
 FT misc\_feature 1..63  
 FT /tag= b  
 FT /label= AcNPV\_polyhedrin\_N-terminal  
 FT 64..1851  
 FT /tag= c  
 FT /label= PT-NANBH\_NS5  
 FT 1852..1875  
 FT /tag= d  
 FT /note= "Synthetic linker region"  
 FT 1876..2706  
 FT /tag= e  
 FT /label= PT-NANBH\_core\_and\_E1\_regions  
 FT 2707..2790  
 FT /tag= f  
 FT /note= "Polyhedrin gene sequence read out-of-frame"  
 FT  
 PN WO9317110-A2.  
 XX  
 PD 02-SEP-1993.  
 XX  
 PP 19-FEB-1993; 93WO-GB000345.  
 XX  
 PR 21-FEB-1992; 92GB-00003803.  
 XX  
 PA (WEL) WELLCOME FOUND LTD.  
 XX  
 PI Parker D, Rodgers BC;  
 XX  
 DR WPI; 1993-288415/36.  
 DR P-PSDB; AAR41431.  
 XX  
 PT New recombinant polypeptide for diagnosing hepatitis C - contains three  
 PT distinct antigens from different viral regions, also useful in protective  
 PT vaccines.  
 PS  
 XX  
 XX Disclosure; Page 20-24; 99pp; English.  
 CC GB-A-2239245 discloses a recombinant polypeptide BHC-11 which comprises  
 CC an antigen obtained from the non-structural coding region (NS) (the 3'  
 CC end) and one antigen from the structural coding region (S) (the 5' end)  
 CC of the NANBH virus. Specifically BHC-11 (AAQ46191) contains a portion of  
 CC the non-structural region of the virus, called NS5, (putative replicase)  
 CC at the N-terminus joined via a synthetic linker to a portion of the  
 CC structural region which contains almost all the core protein sequence (9  
 CC amino acids from the N-terminal are not present) and a part of a sequence  
 CC from the structural region called E1. It is disclosed that BHC-11 may be  
 CC used in diagnosis of PT-NANBH. If at least three different PT-NANBH  
 CC antigens are used to screen for PT-NANBH, the screening is much more  
 CC sensitive as compared to the use of only two PT-NANBH antigens. Pref.  
 CC antigens are described in AAQ46192-94. Two new antigenic regions of the  
 CC PT-NANBH genome are given in AAQ46198-99. AAQ46202 describes an improved  
 CC PT-NANBH recombinant polypeptide. (Updated on 25-MAR-2003 to correct PN  
 CC field.) (Updated on 27-AUG-2003 to correct OS field.)



Query Match 96.5%; Score 804.8; DB 2; Length 3372;  
 Best Local Similarity 99.5%; Pred. No. 5.1e-208;  
 Matches 828; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```

QY 4 AAAACCAACGTAACCAACCACTCCGCCACAGAGAGTACAGGTTCCCGGGCGGTGTGAG 63
DB 2458 AAAACCAACGTAACCAACCACTCCGCCACAGAGAGTACAGGTTCCCGGGCGGTGTGAG 2517
QY 64 ATCGTTGGTAGAGTTTACCTGTGTGCGCGCAGAGGGGCCCAAGGTGGGTGTGCGCGACT 123
DB 2518 ATCGTTGGTAGAGTTTACCTGTGTGCGCGCAGAGGGGCCCAAGGTGGGTGTGCGCGACT 2577
QY 124 AGGAAGACTTCCGAGCGGTGTGCACTCTGTGGAAGCGCAACCTATCCCAAGGCTGCG 183
DB 2578 AGGAAGACTTCCGAGCGGTGTGCACTCTGTGGAAGCGCAACCTATCCCAAGGCTGCG 2637
QY 184 CAGCCCGAGGGGAGGGGCTGAGCCCGGGTACCCCTTGCCCTCATATGCAACGAG 243
DB 2638 CAGCCCGAGGGGAGGGGCTGAGCCCGGGTACCCCTTGCCCTCATATGCAACGAG 2697
QY 244 GGCATGGGGTGGGCAAGATGCTCCTGTACCCCGTGGCTCCCGGCTTAGTTGGGGCCCC 303
DB 2698 GGCATGGGGTGGGCAAGATGCTCCTGTACCCCGTGGCTCCCGGCTTAGTTGGGGCCCC 2757
QY 304 ACTGACCCCGCGGTAGTGTGCGTAAATTTGGGTAAAGTATGATACCTCATATGCGGC 363
DB 2758 ACTGACCCCGCGGTAGTGTGCGTAAATTTGGGTAAAGTATGATACCTCATATGCGGC 2817
QY 364 TTGCGCGACTCTCATATGGGGTACATTTCCGCTGCGCGGCTCCCTTA-GGGGGGCTGCCAG 422
DB 2818 TTGCGCGAC-CTCATATGGGGTACATTTCCGCTGCGCGGCTCCCTTAAGGGGGGCTGCCAG 2876
QY 423 GGCCTTGGCGAGTGGCGTCCGGGTTCTGAGAGACGGCGTGAACCTATGCAACGAGAAATT 482
DB 2877 GGCCTTGGCGAGTGGCGTCCGGGTTCTGAGAGACGGCGTGAACCTATGCAACGAGAAATT 2936
QY 483 ACCCGGTGCTTTTCTCTATCTTCTCTGCTTGGCTTGTCTGTCTGTTGACATTCGACG 542
DB 2937 ACCCGGTGCTTTTCTCTATCTTCTCTGCTTGGCTTGTCTGTCTGTTGACATTCGACG 2996
QY 543 TTCGCTTATGAAGTGGCGCAAGTGTCCGGGATCTACATGACGAACGATGTGCTCAA 602
DB 2997 TTCGCTTATGAAGTGGCGCAAGTGTCCGGGATCTACATGACGAACGATGTGCTCAA 3056
QY 603 CTCAGACATCTGTATGAGACAGCGAGACATGATGACACCCCGGGTGTGTCCCTG 662
DB 3057 CTCAGACATCTGTATGAGACAGCGAGACATGATGACACCCCGGGTGTGTCCCTG 3116
QY 663 TGTCCGGAGGGTAATTCTCCCGCTGCTGGGTAGGCTCACTCCAGGCTTGCGGCCAA 722
DB 3117 TGTCCGGAGGGTAATTCTCCCGCTGCTGGGTAGGCTCACTCCAGGCTTGCGGCCAA 3176
QY 722 GGAAGCGAGCATCCCACTGCGAGCAATAGAGCGCACGTCGATGCTGTGTGGGGCGCG 782
DB 3177 GGAAGCGAGCATCCCACTGCGAGCAATAGAGCGCACGTCGATGCTGTGTGGGGCGCG 3236
QY 783 TSCCTTCTGTCGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 834
DB 3237 TSCCTTCTGTCGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 3288
  
```

RESULT 4  
 AAQ46192  
 ID AAQ46192 standard; cDNA to mRNA; 831 BP.  
 XX  
 AC AAQ46192;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 24-FEB-1994 (first entry)  
 XX  
 DE PT-NANBH virus structural protein region.  
 XX

KM Parenterally transmitted non A non B hepatitis; PT-NANBH;  
 KW hepatitis C virus; HCV; detection; diagnosis; antigen; vaccine; ss.  
 XX  
 OS Hepatitis virus.  
 XX  
 PN W09J17110-A2.  
 XX  
 PD 02-SEP-1993.  
 XX  
 PF 19-FEB-1993; 93WO-GB000345.  
 XX  
 PR 21-FEB-1992; 92GB-00003803.  
 XX  
 PA (WELL) WELLCOME FOUND LTD.  
 XX  
 PI Parker D, Rodgers BC;  
 XX  
 DR WPI; 1993-288415/36.  
 DR P-PSDB; AAR41432.  
 XX  
 PT New recombinant polypeptide for diagnosing hepatitis C - contains three  
 PT distinct antigens from different viral regions, also useful in protective  
 PT vaccines.  
 XX  
 PS Claim 7; Page 29-30; 99PP; English.  
 XX  
 CC If at least three different PT-NANBH antigens are used to screen for PT-  
 CC NANBH, the screening is much more sensitive as compared to the use of  
 CC only two PT-NANBH antigens. Pref. antigens are described in AAQ46192-94.  
 CC Two new antigenic regions of the PT-NANBH genome are given in AAQ46198-  
 CC 99. AAQ46202 describes an improved PT-NANBH recombinant polypeptide.  
 CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to  
 CC correct OS field.)  
 XX  
 SQ Sequence 831 BP; 137 A; 265 C; 251 G; 178 T; 0 U; 0 Other;

Query Match 96.1%; Score 801.6; DB 2; Length 831;  
 Best Local Similarity 99.3%; Pred. No. 2.4e-207;  
 Matches 826; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

QY 4 AAAACCAACGTAACCAACCACTCCGCCACAGAGAGTACAGGTTCCCGGGCGGTGTGAG 63
DB 1 AAAACCAACGTAACCAACCACTCCGCCACAGAGAGTACAGGTTCCCGGGCGGTGTGAG 60
QY 64 ATCGTTGGTAGAGTTTACCTGTGTGCGCGCAGAGGGGCCCAAGGTGGGTGTGCGCGACT 123
DB 61 ATCGTTGGTAGAGTTTACCTGTGTGCGCGCAGAGGGGCCCAAGGTGGGTGTGCGCGACT 120
QY 124 AGGAAGACTTCCGAGCGGTGTGCACTCTGTGGAAGCGCAACCTATCCCAAGGCTGCG 183
DB 121 AGGAAGACTTCCGAGCGGTGTGCACTCTGTGGAAGCGCAACCTATCCCAAGGCTGCG 180
QY 184 CAGCCCGAGGGGAGGGGCTGAGCCCGGGTACCCCTTGCCCTCATATGCAACGAG 243
DB 181 CAGCCCGAGGGGAGGGGCTGAGCCCGGGTACCCCTTGCCCTCATATGCAACGAG 240
QY 244 GGCATGGGGTGGGCAAGATGCTCCTGTACCCCGTGGCTCCCGGCTTAGTTGGGGCCCC 303
DB 241 GGCATGGGGTGGGCAAGATGCTCCTGTACCCCGTGGCTCCCGGCTTAGTTGGGGCCCC 300
QY 304 ACTGACCCCGCGGTAGTGTGCGTAAATTTGGGTAAAGTATGATACCTCATATGCGGC 363
DB 301 ACTGACCCCGCGGTAGTGTGCGTAAATTTGGGTAAAGTATGATACCTCATATGCGGC 360
QY 364 TTGCGCGACTCTCATATGGGGTACATTTCCGCTGCGCGGCTCCCTTA-GGGGGGCTGCCAG 422
DB 361 TTGCGCGAC-CTCATATGGGGTACATTTCCGCTGCGCGGCTCCCTTAAGGGGGGCTGCCAG 419
QY 423 GGCCTTGGCGAGTGGCGTCCGGGTTCTGAGAGACGGCGTGAACCTATGCAACGAGAAATT 482
DB 420 GGCCTTGGCGAGTGGCGTCCGGGTTCTGAGAGACGGCGTGAACCTATGCAACGAGAAATT 479
QY 483 ACCCGGTGCTTTTCTCTATCTTCTCTGCTTGGCTTGTCTGTGTTGACCATTCGACG 542
  
```





KM polymerase chain reaction; diagnostic method; ds.  
 XX Hepatitis C virus.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 325..1734  
 FT /tag= a  
 FT /note= "not full-length CDS, hence no termination codon"  
 PN JP05068562-A.  
 XX  
 XX  
 PD 23-MAR-1993.  
 XX  
 XX  
 PF 30-MAY-1991; 91JP-00153736.  
 XX  
 PR 30-MAY-1991; 91JP-00153736.  
 XX  
 PA (SANWA ) SANWA KAGAKU KENKYUSHO CO.  
 XX  
 XX WPI; 1993-130638/16.  
 DR P-PSDB; AAR34470.  
 XX  
 XX  
 PT DNA and cDNA of hepatitis C virus - useful as probes for diagnosing HCV  
 PT infection.  
 PS Claim 4; Page 20-22; 44pp; Japanese.  
 XX  
 CC CDNA was prepared from HCV genomic RNA. Full-length clone JKI-B (9405  
 CC nucleotides long) and 14 shorter clones were isolated by PCR  
 CC amplification, including clone JK2-A. Primer/probes derived from the  
 CC sequences of these clones can be used in diagnostic assays for HCV. See  
 CC AAQ0425-Q0439  
 XX  
 SQ Sequence 1734 BP; 326 A; 537 C; 505 G; 366 T; 0 U; 0 Other;

Query Match 88.6%; Score 738.6; DB 2; Length 1734;  
 Best Local Similarity 94.5%; Pred. No. 3.9e-190;  
 Matches 787; Conservative 0; Mismatches 44; Indels 2; Gaps 2;

QY 1 AAAAAAAAAAGTAACGTAACCAACCTCCGCCCAAGAGTGAAGTCCCGGGGGTGGT 60  
 DB 349 AAAAAAAAAAGTAACGTAACCAACCTCCGCCCAAGAGTGAAGTCCCGGGGGTGGC 408  
 QY 61 CAGATCGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120  
 DB 409 CAGATCGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 468  
 QY 121 ACTAGGAAGACTTCCGAGCGGTGCGCAACTCGTGAAGGAGCAACTTATCCCAAGGCT 180  
 DB 469 ACTAGGAAGACTTCCGAGCGGTGCGCAACTCGTGAAGGAGCAACTTATCCCAAGGCT 528  
 QY 181 CGCAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
 DB 529 CGCAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 588  
 QY 241 GAGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
 DB 589 GAGGCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 648  
 QY 301 CCCACTGACCCCCCGGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 360  
 DB 649 CCCACTGACCCCCCGGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 708  
 QY 361 GCGTTCGCGAGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 419  
 DB 709 GCGTTCGCGAGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 767  
 QY 420 CAGGCGCTTGGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479  
 DB 768 CAGGCGCTTGGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 827  
 QY 480 TTATCCCGGTTGCTTTCTTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 539

DB 828 TCTGCCGGTGTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 887  
 QY 540 AGCTTCGGCTTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 599  
 DB 888 AGCATCCGCTTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 947  
 QY 600 CAATCAAGCATGTGTAG 659  
 DB 948 CAATCAAGTATTTGTGTAG 1007  
 QY 660 CTGTGTCCGGAGAGGTAATTCCTCCGCTGTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 719  
 DB 1008 CTGTGTCCGGAGAGGTAATTCCTCCGCTGTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1067  
 QY 720 CAGAGCGCAGCATCCCACTGCGACAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779  
 DB 1068 CAGAGCAGCAGCATCCCACTGCGACAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1127  
 QY 780 GCGTCCCTTCTGTGCGCTATGTATGATGATGATGATGATGATGATGATGATGATGATGATG 832  
 DB 1128 GCGTCCCTTCTGTGCGCTATGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1180

## RESULT 7

AAQ0429  
 ID AAQ0429 standard; cDNA, 1734 BP.

AC AAQ0429;  
 XX  
 XX  
 DT 30-JUL-1993 (first entry)  
 XX  
 DE Hepatitis C virus clone JK2-B.  
 XX  
 XX  
 KM HCV; non-A, non-B hepatitis virus; NANBHV; liver disease;  
 KM polymerase chain reaction; diagnostic method; ds.  
 XX  
 FH Hepatitis C virus.  
 OS  
 OS

FH Key Location/Qualifiers  
 FT CDS 325..1734  
 FT /tag= a  
 FT /note= "not full-length CDS, hence no termination codon"  
 FT

PN JP05068562-A.  
 PD 23-MAR-1993.  
 XX  
 XX  
 PF 30-MAY-1991; 91JP-00153736.  
 XX  
 XX  
 PR 30-MAY-1991; 91JP-00153736.  
 XX  
 XX  
 PA (SANWA ) SANWA KAGAKU KENKYUSHO CO.  
 XX  
 XX WPI; 1993-130638/16.  
 DR P-PSDB; AAR34471.  
 XX  
 XX  
 PT DNA and cDNA of hepatitis C virus - useful as probes for diagnosing HCV  
 PT infection.  
 XX  
 XX

PS Claim 4; Page 22-24; 44pp; Japanese.  
 XX  
 CC CDNA was prepared from HCV genomic RNA. Full-length clone JKI-B (9405  
 CC nucleotides long) and 14 shorter clones were isolated by PCR  
 CC amplification, including clone JK2-B. Primer/probes derived from the  
 CC sequences of these clones can be used in diagnostic assays for HCV. See  
 CC AAQ0425-Q0439  
 XX  
 XX

SQ Sequence 1734 BP; 324 A; 540 C; 506 G; 364 T; 0 U; 0 Other;

Query Match 88.6%; Score 738.6; DB 2; Length 1734;  
 Best Local Similarity 94.5%; Pred. No. 3.9e-190;  
 Matches 787; Conservative 0; Mismatches 44; Indels 2; Gaps 2;

QY 1 AGAAAAACAAGTAACCAACCTCCGCCACAGAGCTCAGTTCCCGGCGGTGT 60  
 DB 349 AGAAAAACAAGTAACCAACCTCCGCCACAGAGCTCAGTTCCCGGCGGTGT 408  
 QY 61 CAGATCGTTGGTGAAGTTACCTGTTGCGCGAGAGGCGCCAGTTGGTGTGCGGCG 120  
 DB 409 CAGATCGTTGGTGAAGTTACCTGTTGCGCGAGAGGCGCCAGTTGGTGTGCGGCG 468  
 QY 121 ACTAGGAACATTCGAGCGGTGCGCAACCTCTGTGGAAGGAGCAACCTATCCCCAAGCT 180  
 DB 469 ACTAGGAACATTCGAGCGGTGCGCAACCTCTGTGGAAGGAGCAACCTATCCCCAAGCT 528  
 QY 181 CGCAGACCCGAGGCGAGGCGCTGAGCTCAGCCCGGAGTACCTTGGCCCTCTATAGCAAC 240  
 DB 529 CGCAGACCCGAGGCGAGGCGCTGAGCTCAGCCCGGAGTACCTTGGCCCTCTATAGCAAC 588  
 QY 241 GAGGCGATGGGATGGGAGAGATGGCTCTGTCACCCCGTGGCTCCGCGCTAGTTGGGCG 300  
 DB 589 GAGGCGATGGGATGGGAGAGATGGCTCTGTCACCCCGTGGCTCCGCGCTAGTTGGGCG 648  
 QY 301 CCACTGACCCCGGCGTAACTGCTGCTTAATTTGGGTAAATCATCATACCTTCACATGC 360  
 DB 649 CCACTGACCCCGGCGTAACTGCTGCTTAATTTGGGTAAATCATCATACCTTCACATGC 708  
 QY 361 GGGTTCGCGGACTCTCATGGGATACATTCGCTGTCGCGGCTCCCTTAGG-GGCGCTGC 419  
 DB 709 GGGTTCGCGGACTCTCATGGGATACATTCGCTGTCGCGGCTCCCTTAGG-GGCGCTGC 767  
 QY 420 CAGGCGCTTCGCGCATGGCTCGGCTTCTGAGAGACGCGGTAACTATGCAACAGGAA 479  
 DB 768 CAGGCGCTTCGCGCATGGCTCGGCTTCTGAGAGACGCGGTAACTATGCAACAGGAA 827  
 QY 480 TTATACCGGTTGCTCTTCTCTATCTTCTCTGCTTGGCTTGTGCTGTTGACATTC 539  
 DB 828 TTGACCGGTTGCTCTTCTCTATCTTCTCTGCTTGGCTTGTGCTGTTGACATTC 887  
 QY 540 AGCTTCGCTTATGAAGTGGCAACGATGCGGAGTCTACATGTACAGAACATGCTC 599  
 DB 888 AGCTTCGCTTATGAAGTGGCAACGATGCGGAGTCTACATGTACAGAACATGCTC 947  
 QY 600 CAATCAAGCATGCTGACAGACAGGCAATGATCATGACACCCCGGCTGTGTGCC 659  
 DB 948 CAATCAAGCATGCTGACAGACAGGCAATGATCATGACACCCCGGCTGTGTGCC 1007  
 QY 660 CTGTGTCGCGGAGGATATCTCTCCGCTGCTGGGTAGGCTCACTCCACGCTCGGCG 719  
 DB 1008 CTGTGTCGCGGAGGATATCTCTCCGCTGCTGGGTAGGCTCACTCCACGCTCGGCG 1067  
 QY 720 CAGAGCGCGAGATCCCACTGCGAACATTAAGACGCGCATGATGCTGCTGTTGGGCG 779  
 DB 1068 CAGAGCGCGAGATCCCACTGCGAACATTAAGACGCGCATGATGCTGCTGTTGGGCG 1127  
 QY 780 GGGTTCGCTTCGCTGCGCTATGATGAGTGGGAGATCTCTGCGGATCTGTTTCC 832  
 DB 1128 GGGTTCGCTTCGCTGCGCTATGATGAGTGGGAGATCTCTGCGGATCTGTTTCC 1180

RESULT 8  
 AAQ0430  
 ID AAQ0430 standard; cDNA; 1734 BP.  
 XX  
 AC AAQ0430;  
 XX  
 DT 30-JUL-1993 (first entry)  
 XX  
 DE Hepatitis C virus clone JK2-C.  
 XX  
 KM HCV, non-A, non-B hepatitis virus; NANBHV, liver disease;  
 OS polymerase chain reaction; diagnostic method; ds.  
 XX  
 OS Hepatitis C virus.  
 XX  
 FH Key Location/Qualifiers

FT CDS 325..1734  
 FT /tag= a  
 FT /note= "not full-length CDS, hence no termination codon"  
 XX  
 XX JP05068562-A.  
 XX  
 XX 23-MAR-1993.  
 PD  
 XX  
 XX 30-MAY-1991; 91JP-00153736.  
 PF  
 XX  
 XX 30-MAY-1991; 91JP-00153736.  
 PR  
 XX  
 XX 30-MAY-1991; 91JP-00153736.  
 PA  
 XX  
 XX (SANW) SANWA KAGAKU KENYUSHO CO.  
 DR  
 XX  
 DR WPI; 1993-130638/16.  
 DR P-PSDB; AAR34472.  
 PT  
 XX  
 PT DNA and cDNA of hepatitis C virus - useful as probes for diagnosing HCV  
 PS  
 XX  
 PS Claim 4; Page 24-26; 44p; Japanese.  
 CC  
 XX cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405  
 CC nucleotides long) and 14 shorter clones were isolated by PCR  
 CC amplification, including clone JK2-C. Primer/probes derived from the  
 CC sequences of these clones can be used in diagnostic assays for HCV. See  
 CC AAQ0430-040439  
 XX  
 SQ Sequence 1734 BP; 322 A; 538 C; 508 G; 366 T; 0 U; 0 Other;  
 Query Match 88.4%; Score 737; DB 2; Length 1734;  
 Best Local Similarity 94.4%; Pred. No. 1.1e-189;  
 Matches 786; Conservative 0; Mismatches 45; Indels 2; Gaps 2;  
 QY 1 AGAAAAACAAGTAACCAACCTCCGCCACAGAGCTCAGTTCCCGGCGGTGT 60  
 DB 349 AGAAAAACAAGTAACCAACCTCCGCCACAGAGCTCAGTTCCCGGCGGTGT 408  
 QY 61 CAGATCGTTGGTGAAGTTACCTGTTGCGCGAGAGGCGCCAGTTGGTGTGCGGCG 120  
 DB 409 CAGATCGTTGGTGAAGTTACCTGTTGCGCGAGAGGCGCCAGTTGGTGTGCGGCG 468  
 QY 121 ACTAGGAACATTCGAGCGGTGCGCAACCTCTGTGGAAGGAGCAACCTATCCCCAAGCT 180  
 DB 469 ACTAGGAACATTCGAGCGGTGCGCAACCTCTGTGGAAGGAGCAACCTATCCCCAAGCT 528  
 QY 181 CGCAGACCCGAGGCGAGGCGCTGAGCTCAGCCCGGAGTACCTTGGCCCTCTATAGCAAC 240  
 DB 529 CGCAGACCCGAGGCGAGGCGCTGAGCTCAGCCCGGAGTACCTTGGCCCTCTATAGCAAC 588  
 QY 241 GAGGCGATGGGATGGGAGAGATGGCTCTGTCACCCCGTGGCTCCGCGCTAGTTGGGCG 300  
 DB 589 GAGGCGATGGGATGGGAGAGATGGCTCTGTCACCCCGTGGCTCCGCGCTAGTTGGGCG 648  
 QY 301 CCACTGACCCCGGCGTAACTGCTGCTTAATTTGGGTAAATCATCATACCTTCACATGC 360  
 DB 649 CCACTGACCCCGGCGTAACTGCTGCTTAATTTGGGTAAATCATCATACCTTCACATGC 708  
 QY 420 CAGGCGCTTCGCGCATGGCTCGGCTTCTGAGAGACGCGGTAACTATGCAACAGGAA 479  
 DB 768 CAGGCGCTTCGCGCATGGCTCGGCTTCTGAGAGACGCGGTAACTATGCAACAGGAA 827  
 QY 480 TTATACCGGTTGCTCTTCTCTATCTTCTCTGCTTGGCTTGTGCTGTTGACATTC 539  
 DB 828 TTGACCGGTTGCTCTTCTCTATCTTCTCTGCTTGGCTTGTGCTGTTGACATTC 887  
 QY 540 AGCTTCGCTTATGAAGTGGCAACGATGCGGAGTCTACATGTACAGAACATGCTC 599  
 DB 888 AGCTTCGCTTATGAAGTGGCAACGATGCGGAGTCTACATGTACAGAACATGCTC 947



PR 29-JUL-1994; 94EP-00870132.  
XX  
PA (INNO-) INNOGENETICS NV.  
PI  
PI Maertens G, Bosman F, De Martynoff G, Buyse M;  
DR WPI: 1996-129401/13.  
XX  
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins  
PT - in presence of disulphide bond cleavage agent, to produce proteins  
XX suitable for direct use in vaccines or diagnostic assays of HCV.  
PS  
XX Claim 23; Fig 21; 146pp; English.

AA112704-T12709 and AA112961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques

Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 U; 0 Other;

Query Match	88.4%	Score	737	DB 2:	Length	2433			
Best Local Similarity	94.4%	Pred. NO.	1.2e-189						
Matches	786	Conservative	0	Mismatches	45	Indels	2	Gaps	2

OY	1	AGAAAAACCAACGTAACACAACTCCGCGCCACAGACGTCAAGGTTCCCGGGCGGGGT	60
Db	25	AGAAAAACCAACGTAACACAAACCGCGCCACAGACGTCAAGTTCCCGGGCGGGGT	84
OY	61	CAGATCGTTGGTGGAGTTTACCTGTTCCGCGAGGGGCCAGGTTGGGTGTGCGCG	120
Db	85	CAGATCGTTGGTGGAGTTTACCTGTTCCGCGAGGGGCCAGGTTGGGTGTGCGCG	144
OY	121	ACTAGGAAGACTTCCGAGCGGTTGCAACTCGTGAAGGCGACAACCTATCCCAAGGCT	180
Db	145	ACTAGGAAGACTTCCGAGCGGTTGCAACTCGTGAAGGCGACAACCTATCCCAAGGCT	204
OY	181	CGCCAGCCCGAGGGCGAGGCTCGGAGCTCAAGCCCGGGTACCTTGAGCCCTTAATGGCAAC	240
Db	205	CGCCAGCCCGAGGGGTAGGGCTCGGAGCTCAAGCCCGGGTACCTTGAGCCCTTAATGGCAAT	264
OY	241	GAGGCGATGGGGTGGGCAAGATGGCTCTGTCAACCCCGTGGCTCCGCGCTAAGTTGGGCG	300
Db	265	GAGGCGATGGGGTGGGCAAGATGGCTCTGTCAACCCCGGCGCTCTGGGCTAAGTTGGGCG	324
OY	301	CCCACTGACCCCCGGCGTAAAGTGGGTAAATTTGGGTAAAGTCAATCGTAACTCTCAACATGC	360
Db	325	CCTAAGACCCCCGGCGTAAAGTGGGTAAAGTCAATCGTAACTCTCAACATGC	384
OY	361	GAGCTTCGCGACCTCTCATGGGGTAACTTCCGCTCGTGGGGTCAACCCCTA-GGGGCGGTGC	419
Db	385	GAGCTTCGCGGAC-CTCTGATGGGTAACTTCCGCTCGTGGGGTCAACCCCTAAGGGGGCGGTGC	443
OY	420	CAGGCGCTTGGCGCATGGCGTCCGGGTTCTGAAGACGGCGTGAACCTATGCAACAGGAA	479
Db	444	CAGGCGCTTGGCGCATGGCGTCCGGGTTCTGAAGACGGCGTGAACCTATGCAACAGGAA	503
OY	480	TTTACCAGGTGCTCTTCTCTATCTTCTCTTGGCTTGGCTGTCTGTCTTGAACATTC	539

Db	50	TTTGCCCGATGCTCTTTCTCATCTTCTGCGTTTGCTGTCCTGTCAGACGTTCC	583
Qy	540	AGCTTCGGCTATGAAAGGCGCAAGCTGTCCGGGATCTACCATTTACAGAACTTGCCTC	559
Db	564	AGTTCGCGCTATGAAGTGGCGCAAGCTGTCCGGATGACCATGTACAGAACTGCTCTC	623
Qy	600	CAACTCAAGCATCTGTGTACAGACAGCGGACATGATCATGTACACACCCCGGGTGTGTCC	658
Db	624	CAACTCAAGCATGTGTATGAGGACGAGCATGATCATGTACACACCCCGGGTGTGTCC	683
Qy	660	CTGTGTCCGGGAGGGTAATTCTCCCGCTGCTGGTGTGGCTCATCTCCACGCTCGGGGC	718
Db	684	CTCGTTCGGGAAACCACTCTTCCCGCTGCTGGTGTGGCTCATCTCCACGCTCGGGGC	743
Qy	720	CAAGACGCGACGATCCGCCATCTGCACATACGACGCGATTCGATTGCTGTTGGGAC	779
Db	744	TAGGAACGCGAGGTCCGCCACACGACATACGACGCGATTCGATTGCTGTTGGGAC	803
Qy	780	GGTGGCTTCTGTCGCGATATGTAGTGGGGGATCTGCGGGATCTGTTTTCC	832
Db	804	GGTGGCTTCTGTCGCGATATGTAGTGGGGGATCTGCGGGATCTGTTTTCC	856

RESULT 11  
ADD55557  
ID ADD55557 standard; DNA; 2433 BP.

XX	
DT	15-JAN-2004 (first entry)

DE Hepatitis C virus E1/E2 protein coding sequence #14.

KM Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein.  
KM liver fibrosis; ds; gene.

OS Hepatitis C virus.

PN WO2003051912-A2.

PD 26-JUN-2003.

PF 18-DEC-2002; 2002WO-EP014480.

PR 18-DEC-2001; 2001US-00020510.

XX

2000

XX

DR P-PSDB; ADD55558.

PT New hepatitis C v

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

33

CC Liver disease. Th

CC disease (e.g. liver

XX  
9  
8  
7  
6  
5  
4  
3  
2  
1

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
A  
B  
C  
D  
E  
F  
G  
H  
I  
J  
K  
L  
M  
N  
O  
P  
Q  
R  
S  
T  
U  
V  
W  
X  
Y  
Z

### Best Local Similarity

?

10

```

XX 26-JUN-2003.
PD
XX 18-DEC-2002; 2002WO-EP014480.
XX
XX 18-DEC-2001; 2001US-00020510.
XX
XX 16-OCT-2002; 2002US-0418358P.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Depla E, Bosman F;
XX
XX WPI; 2003-541632/51.
XX
XX P-PSDB; ADD55558.
XX
XX New hepatitis C virus (HCV) vaccine composition, useful for reducing
PT liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.
XX
XX
PS Example 2; SEQ ID NO 49; 271pp; English.
XX
XX The invention comprises an Hepatitis C virus (HCV) vaccine for reducing
CC liver disease. The vaccine of the invention comprises an HCV EI or E2
CC protein as an antigen. The HCV vaccine is useful for reducing liver
CC disease (e.g. liver fibrosis) in a chronic HCV-infected mammal. The
CC present DNA sequence encodes an HCV EI/E2 protein.
XX
XX Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 U; 0 Other;
SQ
Query Match 88.4%; Score 737; DB 10; Length 2433;
Best Local Similarity 94.4%; Pred. NO. 1.2e-189;
Matches 786; Conservative 0; Mismatches 45; Indels 2; Gaps 2
1 AGAAAAACCAAAAGTACCACTCCGCCACAGACGTCCAGCTTCCCGGCGGTGT 60

```

```

Db      25 AGAAAAACCAACGTAACCAACCGCCGCCCAAGAGCTCAAGTTCCGGGGCGGTGCT 84
Qy      61 CAGATCGTTGTGAGAGTTTACCTGTTGCCGCGAGGGGCCCAAGTTGGGTGTGCGCGG 120
Db      85 CAGATCGTTGTGAGAGTTTACCTGTTGCCGCGAGGGGCCCAAGTTGGGTGTGCGCGG 144
Qy      121 ACTAGGAAGACTTCCGAGCGGTGTGCAACTCGTGGAGGCGCAACTTATCCCAAGGCT 180
Db      145 ACTAGGAAGACTTCCGAGCGGTGTGCAACTCGTGGAGGCGCAACTTATCCCAAGGCT 204
Qy      181 CGCGAGCGCGAGGGGCGGGGCTGGGGCTAGCGCCCGGGGTACCTTTGGGCCCTTATGGCAAC 240
Db      205 CGCGAAGCGGAGGAGTGGGCTGGGGCTAGCGCCCGGGGTACCTTTGGGCCCTTATGGCAAT 264
Qy      241 GAGGCGATGGGTGGGCGAGAGTGGCTCTGTCAACCCCGTGGCTCCGGGCTAGTTGGGGC 300
Db      265 GAGGGGATGGGGTGGGCGAGAGTGGCTCTGTCAACCCCGGCGCTCTGGCTAGTTGGGGC 324
Qy      301 CCGACTGACCCCGCGGTAGTGGCTGTAATTGGGTAAGTCAATGATACCTCAGATGC 360
Db      325 CCGACGAGACCCCGGGGTAGTGGCTGTAATTGGGTAAGTCAATGATACCTTATCATGC 384
Qy      361 GGCTTCGCGGACTCTCAATGGGGGTACATCGCGCTGGCGGCTCCCTTA-GGGGGCTGTC 419
Db      385 GGCTTCGCGGAC-CTCGTGGGGTACATTCGCGCTGGCGGCTCCCTTAAGGGGGCTGTC 443
Qy      420 CAGGCGCTGAGCGCATGAGCGTCCGGGTCTGGAGAGCGGCGTGAATATGCAACAGGAA 479
Db      444 CAGGCGCTGAGCGCATGAGCGTCCGGGTCTGGAGAGCGGCGTGAATATGCAACAGGAA 503
Qy      480 TTTTACCGGGTCTCTTTCTTATCTTCTCTTGGCTTGGCTGCTGTTTGAACATTC 539
Db      504 TTTTACCGGGTCTCTTTCTTATCTTCTCTTGGCTTGGCTGCTGTTTGAACATTC 563
Qy      540 AGCTTCGCGTGAAGTGGGCAAGTGTCCGGGATCTACATGTCAGAAAGATTTGTC 599
Db      564 AGCTTCGCGTGAAGTGGGCAAGTGTCCGGGATCTACATGTCAGAAAGATTTGTC 623
Qy      600 CAACTCAAGCATGTGTAGAGAGAGCGGACATGATCATGCAACCCCGGGGTGTGTC 659
Db      624 CAACTCAAGCATGTGTAGAGAGAGCGGACATGATCATGCAACCCCGGGGTGTGTC 683
Qy      660 CTGTGTCCGGAGAGGTAATTCCTCCGCTGTGGTGGTGGCTCACTCCAGCTGTGCGGC 719
Db      684 CTGTGTCCGGAGAGGTAATTCCTCCGCTGTGGTGGTGGCTCACTCCAGCTGTGCGGC 743
Qy      720 CAGAGAGCGGACGATCCGACGTCGACATAGAGAGCGGACGTCGATTTCTGTGGGGC 779
Db      744 TAGAAGCGGACGTCGTCACCAAGCAATACGACGCGCATGATTTCTGTGGGGC 803
Qy      780 GGCTGCTTCTGTGTCGCTATGATGATGAGGGGATCTCTGCGATCTGTTTCC 832
Db      804 GGCTGCTTCTGTGTCGCTATGATGATGAGGGGATCTCTGCGATCTGTTTCC 856

```

RESULT 12  
ADP71139  
ID ADP71139 standard; cDNA; 2433 BP.  
XX

```

AC      ADP71139;
XX
XX      23-SEP-2004 (first entry)
XX
DE      HCV cDNA encoding E1 protein HCC166.
XX
KM      Hepatitis C virus; HCV; E1 glycoprotein; E2 glycoprotein; HCV infection;
KW      liver disease; liver fibrosis; ss; serum alanine aminotransferase level;
XX      steatosis; anti-E2 immunoreactivity; vaccine.
OS      Hepatitis C virus.
XX
XX      US2004126395-A1.
XX

```

```

PD      01-JUL-2004.
XX
XX      18-DEC-2002; 2002US-00321798.
PF
XX
PR      18-DEC-2001; 2001US-0453708P.
PR      16-OCT-2002; 2002US-0418358P.
XX
XX      (MAER/) MAERTENS G.
PA      (DEPL/) DEPLA E.
PA      (BOSM/) BOSMAN F.
PI      Maertens G, Depia E, Bosman F;
XX
XX      WPI; 2004-499089/47.
DR      P-Psdb; ADP71140.
XX
PT      Use of hepatitis C virus (HCV) vaccine composition for reducing liver
PT      disease, serum alanine aminotransferase levels, steatosis, or anti-E2
PT      immunoreactivity in the liver of a chronic HCV-infected mammal.
XX
XX      Disclosure; SEQ ID NO 49; 176bp; English.
XX
CC      The invention relates to the use of a hepatitis C virus (HCV) vaccine
CC      composition for reducing liver disease (such as liver fibrosis or its
CC      progression), serum alanine aminotransferase (ALT) levels, steatosis, or
CC      anti-E2 immunoreactivity in the liver of a chronic HCV-infected mammal,
CC      or for treating a chronic HCV-infected mammal. The liver disease is
CC      reduced by at least 1-2 points according to the overall Ishak score in
CC      the HCV-infected mammal. Also included are a method for predicting
CC      changes in liver disease in a chronic HCV-infected mammal, a therapeutic
CC      HCV vaccine composition (comprising at least one purified or a
CC      combination of at least 2 HCV single or specific oligomeric recombinant
CC      envelope protein selected from an E1 or E2 protein, a part of E1 and E2
CC      proteins, an E1/E2 protein complex formed from purified HCV single or
CC      specific oligomeric recombinant E1 or E2 proteins or its parts and
CC      optionally a pharmaceutical adjuvant), a composition (comprising at least
CC      one E1 or E2 peptide, and optionally, a pharmaceutical adjuvant), an
CC      immunogenic HCV composition (or HCV vaccine composition) comprising a
CC      recombinant virus allowing expression of at least one HCV recombinant
CC      envelope protein (selected from an E1 protein and/or an E2 protein, and
CC      their parts, and optionally, a pharmaceutical adjuvant) and an HCV
CC      vaccine composition (comprising a recombinant virus allowing expression
CC      of at least one HCV recombinant envelope protein chosen from an E1
CC      protein and/or an E2 protein, and parts of the E1 and E2 proteins and,
CC      optionally, a pharmaceutical adjuvant. The HCV vaccine composition is
CC      useful for reducing liver disease (such as liver fibrosis or its
CC      progression), serum ALT levels, steatosis, or anti-E2 immunoreactivity in
CC      the liver in a chronic HCV-infected mammal, or for treating a chronic HCV
CC      -infected mammal, particularly human. The HCV E1 proteins are useful for
CC      in vitro monitoring HCV disease or diagnosing the response to treatment
CC      of patients suffering from HCV infection. The present sequence is an HCV
CC      cDNA encoding an E1 protein (or fragment).
XX
SQ      Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 U; 0 Other;

```

Query Match 88.4%; Score 737; DB 12; Length 2433;  
Best Local Similarity 94.4%; Pred. NO. 1,2e-189;  
Matches 786; Conservative 0; Mismatches 45; Indels 2; Gaps 2;

```

Qy      1 AGAAAAACCAACGTAACCAACCTCGCCCAAGAGCTCAAGTTCCGGGGCGGTGCT 60
Db      25 AGAAAAACCAACGTAACCAACCTCGCCCAAGAGCTCAAGTTCCGGGGCGGTGCT 84
Qy      61 CAGATCGTTGTGAGAGTTTACCTGTTGCCGCGAGGGGCCCAAGTTGGGTGTGCGCGG 120
Db      85 CAGATCGTTGTGAGAGTTTACCTGTTGCCGCGAGGGGCCCAAGTTGGGTGTGCGCGG 144
Qy      121 ACTAGGAAGACTTCCGAGCGGTGTGCAACTCGTGGAGGCGCAACTTATCCCAAGGCT 180
Db      145 ACTAGGAAGACTTCCGAGCGGTGTGCAACTCGTGGAGGCGCAACTTATCCCAAGGCT 204
Qy      181 CGCGAGCGCGAGGGGCGGGGCTGGGGCTAGCGCCCGGGGTACCTTTGGGCCCTTATGGCAAC 240

```

Db 205 CGCCGACCCGAGGGTGAAGGCGCTGAGCGCCGCGGATCCCTTGCCCTCTATGGCAAT 264  
 Qy 241 GAGGCGATGGGGTGGGAGGATGGCTCTCTGCAACCCCGTGGCTCCGCGCTATGTTGGGGC 300  
 Db 265 GAGGCGATGGGGTGGGAGGATGGCTCTCTGCAACCCCGGCGCTCTGGCTTAATTTGGGGC 324  
 Qy 301 CCCACTGACCCCCCGGGGTAGTGGCTGAATTTGGGGTAAATGTCGATACCTCACAATGC 360  
 Db 325 CTTACAGACACCCCGCGGTAGTGGCTGAATTTGGGGTAAATGTCGATACCTCACAATGC 384  
 Qy 361 GGGTTGGCGGACTCTATGAGGGTGAATTCGGCTCTGCGGGCTCCCTTA-GGGGCGCTGC 419  
 Db 385 GGGTTGGCGGAC-CTGGTGGGGTGAATTCGGCTCTGCGGGCGCCCTTAAGGGCGCGCTGC 443  
 Qy 420 CAGGGCGCTGGGCGATGGGCGTCCGGGTTCTGGAGAGCGGGTGAATGCAACAGGGAA 479  
 Db 444 CAGGGCGCTGGGCGATGGGCGTCCGGGTTCTGGAGAGCGGGTGAATGCAACAGGGAA 503  
 Qy 480 TTATCCCGGTGCTCTTCTCTATCTTCTCTTGGCTTGGCTTCTCTGTTGACATTC 539  
 Db 504 TTATCCCGGTGCTCTTCTCTATCTTCTCTTGGCTTGGCTTCTCTGTTGACATTC 563  
 Qy 540 AGCTTCCGCTTATGAAGTGGCGCAAGTTCGGGATCTACATGTCAGCAAGATTGCTC 599  
 Db 564 AGCTTCCGCTTATGAAGTGGCGCAAGTTCGGGATCTACATGTCAGCAAGATTGCTC 623  
 Qy 600 CAATCAGAGCATGTATGAGAGCAGGCGCATGATCATGCAACCCCGGGGTGGTGC 659  
 Db 624 CAATCAGAGCATGTATGAGAGCAGGCGCATGATCATGCAACCCCGGGGTGGTGC 683  
 Qy 660 CTGTGTCCGGAGGATTAATCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCCGCGC 719  
 Db 684 CTGTGTCCGGAGAAACAATCTTCCGCTGCTGGGTAGCGCTCACTCCACGCTCCGCGC 743  
 Qy 720 CAAGACGCCAGCATATCCCACTCCGCAATGACAGCGCACATGTTGCTGTTGGGAC 779  
 Db 744 TGGAGACGCCAGCATATCCCACTCCGCAATGACAGCGCACATGTTGCTGTTGGGAC 803  
 Qy 780 GGGTGGCTTCTGTGCTGCTATGAGCGGGGAGTCTCTGGGATCTGTTTTC 832  
 Db 804 GGGTGGCTTCTGTGCTGCTATGAGCGGGGAGTCTCTGGGATCTGTTTTC 856  
 RESULT 13  
 ADR82189  
 ID ADR82189 standard; DNA; 9587 BP.  
 AC ADR82189;  
 DT 16-DEC-2004 (first entry)  
 XX Hepatitis C virus type 1b polypeptide DNA.  
 DE Hepatitis C virus type 1b polypeptide DNA.  
 XX antilipemic; cardiant; vasotropic; antiarteriosclerotic; antidiabetic;  
 XX cytostatic; anticonvulsant; nootropic; muscular; anti-HIV;  
 XX RNA interference; iRNA; antisense technology; lipid metabolism;  
 XX cholesterol imbalance; dyslipidaemia hypercholesterolaemia;  
 XX coronary artery disease; CAD; coronary heart disease; CHD;  
 XX atherosclerosis; hepatic glucose production;  
 XX glucose-metabolism-related disorder; diabetes; cancer; breast cancer;  
 XX colon cancer; lung cancer; neurological disease; Huntington disease;  
 XX spinocerebellar ataxia; viral disease; AIDS; HCV, polypeptide; gene; ds.  
 OS Hepatitis C virus.  
 XX  
 XX WO2004080406-A2.  
 PD 23-SEP-2004.  
 XX  
 PF 08-MAR-2004; 2004MO-US007070.  
 XX  
 PF 07-MAR-2003; 2003US-0452682P.  
 PR 12-MAR-2003; 2003US-0454265P.

PR 13-MAR-2003; 2003US-0454962P.  
 PR 13-MAR-2003; 2003US-0455050P.  
 PR 14-APR-2003; 2003US-0462894P.  
 PR 17-APR-2003; 2003US-0463772P.  
 PR 25-APR-2003; 2003US-0465665P.  
 PR 25-APR-2003; 2003US-0465802P.  
 PR 09-MAY-2003; 2003US-0469612P.  
 PR 08-AUG-2003; 2003US-0493986P.  
 PR 11-AUG-2003; 2003US-0494597P.  
 PR 26-SEP-2003; 2003US-0506341P.  
 PR 09-OCT-2003; 2003US-0510246P.  
 PR 10-OCT-2003; 2003US-0510318P.  
 PR 07-NOV-2003; 2003US-0518453P.  
 XX  
 PA (ALNY-) ALNYLAM PHARM.  
 XX  
 XX Manoharan M, Bumcrot D;  
 XX MPI; 2004-677362/66.  
 DR  
 XX  
 PT Interference RNA agent useful for treating dyslipidemias, coronary artery  
 PT disease, diabetes, cancer or neurological disease, comprises sense  
 PT sequence and antisense sequence which has specific modifications.  
 XX  
 PS Example 5; SEQ ID NO 6688; 378bp; English.  
 XX  
 CC The invention describes a RNA interference (iRNA) agent (I) comprising a  
 CC sense sequence and an antisense sequence, where the sense sequences have  
 CC one or more asymmetrical 2'-O-alkyl modifications, the antisense  
 CC sequences have one or more asymmetrical phosphorothioate modifications  
 CC and the antisense sequence targets a human gene sequence. Also described  
 CC are: a pharmaceutical preparation comprising (I); reducing (M1) apob-100  
 CC levels or glucose-6-phosphatase levels in a subject; producing (I);  
 CC stabilising (I), involves selecting a sequence with activity and  
 CC introducing one or more asymmetrical modification in the sequence, where  
 CC the modification decreases nuclease sensitivity while not decreasing its  
 CC activity; a kit comprising (I) and an instruction for its use; and a device  
 CC that can be dispense or administer a composition comprising (I). (I) is  
 CC useful for reducing apob-100 levels or glucose-6-phosphatase levels. (M1)  
 CC is useful for reducing apob-100 levels or glucose-6-phosphatase levels.  
 CC The subject is suffering from a disorder characterised by elevated or  
 CC otherwise unwanted expression of apob-100, elevated or otherwise unwanted  
 CC levels of cholesterol, and/or dysregulation of lipid metabolism. The  
 CC disorder is chosen from the HDL/LDL cholesterol imbalance,  
 CC dyslipidaemia, hypercholesterolaemia, statin-resistant  
 CC hypercholesterolaemia, coronary artery disease (CAD), coronary heart  
 CC disease (CHD) and atherosclerosis. (I) is administered to a subject to  
 CC inhibit hepatic glucose production or for treating glucose-metabolism-  
 CC related disorder e.g. diabetes or type-2 diabetes. (I) is useful for  
 CC treating the diseases as mentioned above, cancer (e.g. breast, colon or  
 CC lung cancer), neurological disease (e.g., Huntington disease or  
 CC spinocerebellar ataxia) or viral disease (e.g., AIDS). This sequence  
 CC represents hepatitis C virus type 1b polypeptide DNA.  
 XX  
 SQ Sequence 9587 BP; 1921 A; 2865 C; 2714 G; 2087 T; 0 U; 0 Other;  
 Query Match 88.4%; Score 737; DB 13; Length 9587;  
 Best Local Similarity 94.4%; Pred. No. 1.8e-189;  
 Matches 786; Conservative 0; Mismatches 45; Indels 2; Gaps 2;  
 Qy 1 AGAAAAACAACGTAACCAACCTCCGCGCAGAGAGTACGTTCCCGGCGGTGGT 60  
 Db 366 AGAAAAACAACGTAACCAACCTCCGCGCAGAGAGTACGTTCCCGGCGGTGGT 425  
 Qy 61 CAGATGTTGTGAGATTACCTGTTCCGCGCAGAGGCGCCAGGTTGGGTGCGCGG 120  
 Db 426 CAGATGTTGTGAGATTACCTGTTCCGCGCAGAGGCGCCAGGTTGGGTGCGCGG 485  
 Qy 121 ACTAGGAACATTTCCGAGCGGTGCGCAACTGTGTGAAGGCGCAACTTATCCCAAGGT 180  
 Db 486 ACTAGGAACATTTCCGAGCGGTGCGCAACTGTGTGAAGGCGCAACTTATCCCAAGGT 545  
 Qy 181 CGCAGCGCGAGGCGAGGCGCTGGGCTACGCCCGGTTACCTTGGCCCTTATGGCAAC 240





QY 780 GGCTGCTTCTCCGCTATGATGAGGAGATCTCTCGGATCTGTTCC 832  
 DB 804 GGCTGCTTCTGCTCCGCTATGATGAGGAGATCTCTCGGATCTGTTCC 856

## RESULT 15

AAQ43889 standard; cDNA to mRNA; 2540 BP.

AAQ43889;

21-OCT-1993 (first entry)

NANB hepatitis virus polynucleotide N-2540-2.

Non-A, non-B; virus; polymerase chain reaction; detection; sensitive; specific; HCV; NANBH; ss.

non-B hepatitis virus.

Location/Qualifiers

1.341

/\*tag= b

/note= "From 5' terminal of NANBH virus RNA"

342.2540

/\*tag= a

JP05091884-A.

16-APR-1993.

10-APR-1991; 91JP-00196175.

12-JUN-1990; 90JP-00153401.

08-NOV-1990; 90JP-00304405.

(NAKA/) NAKAMURA T.

WPI; 1993-199637/25.

P-PSDB; AAR38279.

Antigen related to non-A and non-B hepatitis virus - comprises non-

translation region comprising 340 - 341 mols. of nucleotides, non-

translation region comprising 1885 - 2551 mols. of nucleotides including

region 1,149 and, etc.

Claim 3; Page 19-20; 73JP; Japanese.

The sequence is that of NANB hepatitis virus polynucleotide N-2540-2

which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM. The

polypeptide it encodes may be used in a system for detecting NANB

hepatitis. This method is highly specific and sensitive, and can detect

NANB hepatitis virus which could not be detected by conventional methods

Sequence 2540 BP; 471 A; 775 G; 741 G; 553 T; 0 U; 0 Other;

Query Match 88.0%; Score 733.8; DB 2; Length 2540;

Best Local Similarity 94.1%; Pred. No. 8.8e-189;

Matches 784; Conservative 0; Mismatches 47; Indels 2; Gaps 2;

1 AGAAAAACCAAAAGCTAACCAACCTCCGCCACAGAGCTCAAGTTCCCGGCGGTGCT 60

366 AGAAAAACCAAAAGCTAACCAACCTCCGCCACAGAGCTCAAGTTCCCGGCGGTGCT 425

61 CAGATCGTGTGAGTGAAGTTTACCTGTTGCCGCGAGGAGGCCCAAGTTGGTGTGCGCG 120

426 CAGATCGTGTGAGTGAAGTTTACCTGTTGCCGCGAGGAGGCCCAAGTTGGTGTGCGCG 485

121 ACTAGGAAGACTTCCGAGCGGTGCAACCTGTGGAAGGCGACACCTATCCCAAAGCT 180

486 ACTAGGAAGACTTCCGAGCGGTGCAACCTGTGGAAGGCGACACCTATCCCAAAGCT 545

QY 181 CGCAGACCCGAGGAGGAGGCTGAGGCTCAGCCCGGATACCTTGGCCCTTATGCAAC 240  
 DB 546 CGCCAGACCCGAGGAGGAGGCTGAGGCTCAGCCCGGATACCTTGGCCCTTATGCAAC 605  
 QY 241 GAGGAGTGGGAGGAGGAGGCTGAGGCTCAGCCCGGATACCTTGGCCCTTATGCAAC 300  
 DB 606 GAGGAGTGGGAGGAGGAGGCTGAGGCTCAGCCCGGATACCTTGGCCCTTATGCAAC 665  
 QY 301 CCCACTGAGACCCCGGAGGAGGCTGAGGCTCAGCCCGGATACCTTGGCCCTTATGCAAC 360  
 DB 666 CCCACTGAGACCCCGGAGGAGGCTGAGGCTCAGCCCGGATACCTTGGCCCTTATGCAAC 725  
 QY 361 GGCTTGGCCGACTGATGAGGAGTACATTCGCTGCTGAGGCTGCTTGA-GGAGCGCTGC 419  
 DB 726 GGCTTGGCCGACTGATGAGGAGTACATTCGCTGCTGAGGCTGCTTGA-GGAGCGCTGC 784  
 QY 420 CAGGAGCTTGGCAGACAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 479  
 DB 785 CAGGAGCTTGGCAGACAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 844  
 QY 480 TTACCCGGTGTGCTTCTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 539  
 DB 845 CTGCTCCGGTGTGCTTCTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 904  
 QY 540 AGCTTCGCTTATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 599  
 DB 905 AGCTTCGCTTATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 964  
 QY 600 CAATCAAGCATGTGTATGAG 659  
 DB 965 CAATCAAGCATGTGTATGAG 1024  
 QY 660 CTGTGTCGGGAGAGGAGTATCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 719  
 DB 1025 CTGTGTCGGGAG 1084  
 QY 720 CAGAGAGCAGAGATCCCACTGCGACCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779  
 DB 1085 CAGAGAGCAGAGATCCCACTGCGACCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1144  
 QY 780 GGCTGCTTCTGCTCGCTATGATGAGTGGGAGATCTTGGAGATCTGTTTC 832  
 DB 1145 GGCTGCTTCTGCTCGCTATGATGAGTGGGAGATCTTGGAGATCTGTTTC 1197

Search completed: February 21, 2005, 01:56:31  
 Job time : 595 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 01:13:20 : Search time 3559 Seconds  
(without alignments)  
8919.809 Million cell updates/sec

Title: US-09-664-363-5

Perfect score: 834  
Sequence: 1 AGAAAAACCAAGTAACAC.....TTCGCGATCTGTTTCCG 834

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57.4	6.9	1100	CNS016KD	AL106855 Drosophila
2	56.2	6.7	925	CNS0091P	AL055013 Drosophila
3	53.4	6.4	879	CNS020QG	AL059545 Tetradon
4	48.6	5.8	1101	CNS016UE	AL107216 Drosophila
5	48	5.8	1101	CNS008NG	AL057398 Drosophila
6	47.8	5.7	1036	CNS010BS	AL098470 Drosophila
7	47.6	5.7	1030	AG126333	AG126333 Pan trogl
8	45.8	5.5	925	CNS0091P	AL055013 Drosophila
9	44.8	5.4	844	CNS0052P	AL056652 Drosophila
10	44.8	5.4	910	CNS0060N	AL066629 Drosophila
11	44.4	5.3	439	CO141650	CO141650 EST836321
12	44.4	5.3	450	CO137934	CO137934 EST836305
13	44.4	5.3	743	CO138520	CO138520 EST833191
14	43.6	5.2	548	CO143444	CO143444 EST838115
15	43.2	5.2	899	CD109759	CD109759 AGENCOURT
16	43.2	5.2	932	CNS00720	AL066742 Drosophila
17	43	5.2	1201	CNS016BR	AL106545 Drosophila
18	42.6	5.1	932	CNS00720	AL066742 Drosophila
19	42.6	5.1	936	BG853371	BG853371 102403440
20	42	5.0	400	CO137412	CO137412 EST832083
21	41.8	5.0	442	AF367693	AF367693 AF367693
22	41.6	5.0	692	CNS007WH	AL050923 Drosophila
23	41.6	5.0	839	CNS004NB	AL054280 Drosophila
24	41.6	5.0	1135	BUS27635	BUS27635 AGENCOURT

25	41.2	4.9	387	7	CO148295	CO148295 EST823348
26	41.2	4.9	1009	9	CNS010EW	AL098882 Drosophila
27	40.8	4.9	518	9	CNS016UP	AL107227 Drosophila
28	40.8	4.9	997	9	CNS006DN	AL065132 Drosophila
29	40.6	4.9	980	9	AG071642	AG071642 Pan trogl
30	40.4	4.8	1203	9	CNS015Y4	AL106054 Drosophila
31	40.4	4.8	1233	6	CA975828	CA975828 AGENCOURT
32	40.2	4.8	405	7	CO893216	CO893216 Bowen 21
33	40.2	4.8	466	2	AM670394	AM670394 114401_MA
34	40.2	4.8	548	7	CK952578	CK952578 4091981_B
35	40.2	4.8	781	7	CO876886	CO876886 Bowen_05
36	40.2	4.8	801	7	CK847989	CK847989 970689_MA
37	40	4.8	842	9	AG058791	AG058791 Pan trogl
38	40	4.8	933	5	BX383247	BX383247 BX383247
39	39.6	4.7	1014	6	CB204069	CB204069 AGENCOURT
40	39.4	4.7	356	5	BY235468	BY235468 BY235468
41	39.4	4.7	750	9	AG425278	AG425278 Mus muscu
42	39.4	4.7	798	5	BP144487	BP144487 BP144487
43	39.4	4.7	869	9	AG162064	AG162064 Pan trogl
44	39.4	4.7	1101	9	CNS0178Y	AL108460 Drosophila
45	39.2	4.7	761	9	CG373319	CG373319 OGB144TV

## ALIGNMENTS

RESULT 1  
CNS016KD  
LOCUS  
DEFINITION  
CNS016KD Drosophila melanogaster genome survey sequence Sp6 end of BAC  
BACN16D22 of Drosophila library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AL106855  
AL106855.1 GI:5624152  
GSS.  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Bukayeva; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Phyloidea; Drosophilidae; Drosophila.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Genoscope.  
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billand at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

FEATURES  
source  
1..1100  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACN16D22"  
/clone\_1lb="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : Sp6"

## ORIGIN

Query Match 6.9%; Score 57.4; DB 9; Length 1100;  
Best local Similarity 21.9%; Pred. No. 0.00045;  
Matches 114; Conservative 111; Mismatches 236; Indels 0; Gaps 0;  
25 CTCGCCACAGAGCTCAGGTTCCCGGCGGTGTCAGATCGTGTGAGTTTACCTG 84  
DB 556 CGCTGTCMGTGCTGKKGKGTGGCGKSGTSGGSGKGTGCTGTGCTGGCGCCTGG 615

Oy 85 TTGCGCGACGAGGGGCCCCAGGTTGGGTGTGGCGCGAATTAGAAAGCTTTCGAGCGTCG 144  
 Db 616 TGTATTKSGKGYGGCCBCGSGSGGCKSTTCGSGSGGGGAGYKSGKSGSGCCGCGSGCGSGSG 675  
 Oy 145 CAACCTCGTGAAGCGCAACAACCTATCCCAAGGCTCGACAGCCCGAGAGGAGGCGCTTG 204  
 Db 676 CGCSGSSSGCGKCGSKSGSCSBBSBGYCCBGSBGYCKCSBGCGGCGGCGCGGSGC 735  
 Oy 205 GCTAGAGCCCGGGTACCCCTTGGCCCTCTATATGGCAACAGAGGCGATGGGGTGGGACAGATGG 264  
 Db 736 CGCGTCGGCKGBBGGGCGYGYCGGSKCCBCTGKTCCSCYSTCKCKSGCGGTGTGTCCK 795  
 Oy 265 TTCTGTGACCCCGTGGCTCCCGGCTATGTTGGGGCCCCACTGACCCCGGCGTAGGTGCG 324  
 Db 796 CKCTGTGCKBKYCKCYCTTCKYBCKBCKYCKCGCGGCGSBKSGCGGCKCKTCKKCY 855  
 Oy 325 CGTAAATTTGGGTAAGATCATGATGAACCTCAACATGCGGGCTTGGCCGACTCTCATAGGGGTA 384  
 Db 856 KKKKKRYBKRYKKCKCTYKCKCKCBCKCKYKCKBKCKCKBKCKCKCKCKCKCKCKBK 915  
 Oy 385 CATTCCGCTCGTGGCGCGCTCCCTTAAGAGGGCGGTGCAAGGGCGTGGGCGCATAGGGGTCCGG 444  
 Db 916 BKCBKCBKCKCBKCBKCCCBKBYKCCCTCBKCBKCKYKCKCKCKBKBYKCBKCBKCK 975  
 Oy 445 GTTCTGAGAGCGGGTGAACATATGACAAGAGGAATTTACCGGGTGTCTTTCTATC 504  
 Db 976 KCCCBKCBKCKCKCBKCKCKKCKKCKBKCBKCBKCBKCKCKCKCKCKCKCKCKKYYT 1035  
 Oy 505 TTCTCTTTGGCTTGCTGTCTCTGTTTACCATTCAGGTTTC 545  
 Db 1036 YTBKKYKRYKKRYKRYKTBKKRYKBYKCKCYKCBKSBKCKCKBKCK 1076

LOCUS	CNS009J1P	925 bp	DNA	linear	GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence Tetr3 end of RCP1 #				
ACCESSION	BACR19D16 of RCP1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
VERSION	AL053013				
KEYWORDS	AL053013.1	GI:4934461			
SOURCE	GSS.				
ORGANISM	Drosophila melanogaster (fruit fly)				
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peiryogeta; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridroidea; Drosophilidae; Drosophila. 1 (bases 1 to 925) Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org/The_BDGP_Drosophila_melanogaster_BAC_library">http://www.fruitfly.org/The_BDGP_Drosophila_melanogaster_BAC_library</a> was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RCP1-98 and was constructed by partial ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .				
COMMENT					

FEATURES	Location/Qualifiers
source	1..925
	/organism="Drosophila melanogaster"
	/mol_type="genomic DNA"
	/db_xref="taxon:7227"

```

/clone="BACR19D16"
/clone_lib="RPCT-98"
/note="end : TET3"

```

[illegible]

RESULT 3	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	
CNS0200G	879 bp	DNA	linear	GSS 01-SEP-2000	Tetradon nigroviridis genome survey sequence T7 end of clone 158C06 of library G from Tetradon nigroviridis, genomic survey sequence.	AL209545	GI:7868364	GSS; genome survey sequence.	Tetradon nigroviridis	Tetradon nigroviridis	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetracodontiformes; Tetracodontidae; Tetracodontidae; Tetracodon.	1	Roest Crolius,H., Jalllon,O., Dasliya,C., Bouneau,L., Fisher,C., Bernic,A., Fzames,C., Wincker,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.	Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence	Nat. Genet. 25 (2), 235-238 (2000)	10835645	2	Roest Crolius,H., Jalllon,O., Dasliya,C., Ozouf-Costaz,C., Fzames,C., Fischer,C., Bouneau,L., Billaud,A., Quetier,F., Saurin,W., Bernic,A. and Weissenbach,J.	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis	Genome Res. 10 (7), 939-949 (2000)	20359837	3	(bases 1 to 879)	Genoscope.	Direct Submission	Submitted (12-ARR-2000) Genoscope - Centre National de Sequençage BP 191 91006 Evry cedex - FRANCE (E-mail : segrefigenoscope.cns.fr - web : www.genoscope.cns.fr)

**Source**

The BDGP is constructing a physical map of the *Drosophila* melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila\\_melanogaster\\_BAC\\_library](http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library) was prepared by Kazutoyo Osoegawa and Aaron Mammoso in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCR-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).



Vector : PK145  
R.Site 1 : SacI  
R.Site 2 : SacI  
Location/Qualifiers  
1. .1030  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="PTB-136M19.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC library"

## ORIGIN

Query Match 5.7%; Score 47.6; DB 9; Length 1030;  
Best Local Similarity 34.3%; Pred. No. 0.17;  
Matches 134; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

25 CTCGCCACACAGAGACGTCCGAGGTCGTCAGATCGTTGAGTTTACTCTG 84  
480 CCGCACACACCCCGCCGCGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 421  
85 TTGCCCGCAGAGGCGCCCAAGTTGCTGCGCGCACTAGAAAGACTTCGACGCTG 144  
420 GGGCGGGGGGGGGCGCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 361  
145 CAACCTCTGTGAAGGAGCAACTATCCCAAGGCTCGCCAGCGGAGGCGCTTG 204  
360 GGGCGCGGAGGAGGCGCGCAACGACCCCGCGCGCGGGGGGGGGGGGGGG 301  
205 GCTCAGCCCGGAGTACCTTGAGCCCTCTATGCAAGAGGAGGATGGGAGTGG 264  
300 GGGGGAGNNNGGAGGGGGGGGGGCGGAGGAGGAGGGGGGGGGGGAGNG 241  
265 CTCCTGTACACCCGCTGCTCCGCGCTTATGGGAGCCCACTAGCCCGCTAGTGC 324  
240 NNNNNNGAGNN 181  
325 CGTATTGGGTAAAGTCAATCATACCTTCATCGGCGCTTCCCGACTTCATGGAGTA 384  
180 NGAGCNCNN 121  
385 CATTCGCTGCTGCGGCTCCCTTAGGGGCG 415  
120 CANNCCNN 90

## RESULT 8

CNS0091P/c

925 bp DNA linear GSS 03-JUN-1999

LOCUS  
DEFINITION  
BACR19D16 of RPCI-98 library from *Drosophila melanogaster* (fruit fly), genomic survey sequence.

AL053013 GI:4934461

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
GSS.  
*Drosophila melanogaster* (fruit fly)  
*Drosophila melanogaster*  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; *Drosophila*.  
1 (bases 1 to 925)

## REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

Direct Submision  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley *Drosophila* Genome Project (BDGP).  
The BDGP is constructing a physical map of the *Drosophila*  
*melanogaster* genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP *Drosophila*  
*melanogaster* BAC library was prepared by Kazutoyo Osoegawa and

## FEATURES

source

Location/Qualifiers  
1. .925  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR19D16"  
/clone\_lib="RPCI-98"  
/note="end : TET3"

## ORIGIN

Query Match 5.5%; Score 45.8; DB 9; Length 925;  
Best Local Similarity 12.1%; Pred. No. 0.49;  
Matches 43; Conservative 163; Mismatches 149; Indels 0; Gaps 0;

88 CCGCGAGGGGCCCCAGATTGGGTGCGCGGACTAGAGACTTCCGAGCGTGCAG 147  
924 SBSGSCSCSSBSSCSCSSSMSTSSBSBSSCSCSSBSSSSTSSBSSBSSSS 865  
148 CCTGTGAGAGGAGACACTTATCCCAAGCTCGCCAGCCGAGGAGCGGCTGCT 207  
864 SGTSSACVACNMSBSCCGCGMACBMCWSSSSSCGASABRGVAKRAGAGKRGGSG 805  
208 CAGCCCGGATACCTTGCGCCCTCTATGCAAGAGGAGGATGGGAGGAGATGGC 267  
804 GASASSSSSACBSSSSCSASCSWSSSSSSSRGAGAGGAGSSSSSSSSA 745  
268 CTGTACACCCGCTGCTCCGCGCTAGTGGGCGCCCACTGAGCCCGCGTATGCTG 327  
744 SAGSVSSASSSSSSCSSSVSSCSVASMSBSSSSASBSSSSSSASCSGCCCT 685  
328 AATTGGGTAAAGTCAATCATACCTTCATCGGCGCTTCCCGACTTCATGGAGTA 387  
684 SWSGSCSTASMSAARSSSSSSSCSSSMSASSSASSSSSSSSSSGSACGS 625  
388 TCCGCTGCTGCGGCTCCCTTAGGGGCGCTGCCAGGCGCCCTGGCGCATGGCTCC 442  
624 MSSGGGSGSVASBSSGMSVSSSGRSGSGGAGVGGSSSSSGSGSGGSGV 570

## RESULT 9

CNS0052P/c

844 bp DNA linear GSS 03-JUN-1999

LOCUS  
DEFINITION  
BACR11P16 of RPCI-98 library from *Drosophila melanogaster* (fruit fly), genomic survey sequence.

AL056652 GI:4932342

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
GSS.  
*Drosophila melanogaster* (fruit fly)  
*Drosophila melanogaster*  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; *Drosophila*.  
1 (bases 1 to 844)

## REFERENCE

AUTHORS

TITLE

## COMMENT

Direct Submision  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley *Drosophila* Genome Project (BDGP).  
The BDGP is constructing a physical map of the *Drosophila*  
*melanogaster* genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP *Drosophila*





FORWARD: M13F  
 BACKWARD: M13R  
 Seq primer: M13 Forward  
 POLYA=No.

FEATURES  
 Source Location/Qualifiers

1. 439  
 /organism="Aspergillus flavus"  
 /mol\_type="mRNA"  
 /strain="NRRL 3357"  
 /db\_xref="taxon:5059"  
 /clone="NAFE03"  
 /sex="asexual mycelia"  
 /cell\_type="mycelia"  
 /dev\_stage="developmental stages from 18 to 96 hours"  
 /lab\_host="E. coli DH10B T1 resistant cells"  
 /clone\_1ib="Aspergillus flavus Normalized cDNA Expression Library"  
 /note="Vector: pBluescript (SK+) (Stratagene), antibiotic selection marker: Carbenicillin; Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

## ORIGIN

Query Match 5.3%; Score 44.4; DB 7; Length 439;  
 Best Local Similarity 49.2%; Pred. No. 1;  
 Matches 117; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

579 CCATGTCAGCAAGATGCTCCAACTCAAGCATGTGTAGAGACAGCGGACATGATCAT 638  
 32 CAAGCGCGCGCCCAAGACACCTCTTAAGATCCAGATGTTGTCTCCAGACCAACA 91  
 639 GCACACCGCGGGGTGTGTGCTGTGTCCGAGAGGTAATCTCCGCTGTGGTAGC 698  
 92 GCTGCGCCCCAAGGGGAGTTGGGTCGTGTGATCTGGGCTACTGCTGGGGTGA 151  
 699 GCTCACTCCACGCTGCGCGCAAGAGCCGACATCCCACTGCGCAATACAGCGCA 758  
 152 GATCACCGCGCTGCGCTCGGGCAAGGACAAAGATCTTCAAGCGCGCTGCAATG 211  
 759 CGTCGATTTGCTGCTGTGGGGGCGGCTCTGCTGCTATGATGATGAGTGGGATCTC 816  
 212 CCCCGTATGCTTGTCCCAAGATGGAAGTCGCTCAATCCCGCTGCTC 269

RESULT 12 450 bp mRNA linear EST 17-JUN-2004  
 LOCUS CO137934  
 DEFINITION Aspergillus flavus Normalized cDNA Expression Library  
 ACCESSION CO137934  
 VERSION CO137934.1 GI:48886912  
 KEYWORDS EST.  
 SOURCE Aspergillus flavus  
 ORGANISM Aspergillus flavus  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 1 (bases 1 to 450)  
 Yu, J., Whitelaw, C. A., Nierman, W. C., Bhatnagar, D. and Cleveland, T. E.  
 Aspergillus flavus expressed sequence tags for identification of genes with putative roles in aflatoxin contamination of crops  
 FEMS Microbiol. Lett. (2004) In press  
 CONTACT: Yu J  
 JOURNAL Food and Feed Safety Research Unit  
 USDA/ARS, Southern Regional Research Center  
 1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA  
 Tel: 504 286 4405  
 Fax: 504 286 4419  
 Email: jiyu@ars.ars.usda.gov  
 Contact Dr. Yu at USDA/ARS SRRC (jiyu@ars.ars.usda.gov) for clone

Information  
 PCR Primers  
 FORWARD: M13F  
 BACKWARD: M13R  
 Seq primer: M13 Forward  
 POLYA=No.

FEATURES  
 Source Location/Qualifiers

1. 450  
 /organism="Aspergillus flavus"  
 /mol\_type="mRNA"  
 /strain="NRRL 3357"  
 /db\_xref="taxon:5059"  
 /clone="NAFCN29"  
 /sex="asexual mycelia"  
 /cell\_type="mycelia"  
 /dev\_stage="developmental stages from 18 to 96 hours"  
 /lab\_host="E. coli DH10B T1 resistant cells"  
 /clone\_1ib="Aspergillus flavus Normalized cDNA Expression Library"  
 /note="Vector: pBluescript (SK+) (Stratagene), antibiotic selection marker: Carbenicillin; Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

## ORIGIN

Query Match 5.3%; Score 44.4; DB 7; Length 450;  
 Best Local Similarity 49.2%; Pred. No. 1;  
 Matches 117; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

579 CCATGTCAGCAAGATGCTCCAACTCAAGCATGTGTAGAGACAGCGGACATGATCAT 638  
 43 CAAGCGCGCGCCCAAGACACCTCTTAAGATCCAGATGTTGTCTCCAGACCAACA 102  
 639 GCACACCGCGGGGTGTGTGCTGTGTCCGAGAGGTAATCTCCGCTGTGGTAGC 698  
 103 GCTGCGCCCCAAGGGGAGTTGGGTCGTGTGATCTGGGCTACTGCTGGGGTGA 162  
 699 GCTCACTCCACGCTGCGCGCAAGAGCCGACATCCCACTGCGCAATACAGCGCA 758  
 163 GATCACCGCGCTGCGCTCGGGCAAGGACAAAGATCTTCAAGCGCGCTGCAATG 222  
 759 CGTCGATTTGCTGCTGTGGGGGCGGCTCTGCTGCTATGATGATGAGTGGGATCTC 816  
 223 CCCCGTATGCTTGTCCCAAGATGGAAGTCGCTCAATCCCGCTGCTC 280

RESULT 13 743 bp mRNA linear EST 17-JUN-2004  
 LOCUS CO138520  
 DEFINITION Aspergillus flavus Normalized cDNA Expression Library  
 ACCESSION CO138520  
 VERSION CO138520.1 GI:48887498  
 KEYWORDS EST.  
 SOURCE Aspergillus flavus  
 ORGANISM Aspergillus flavus  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 1 (bases 1 to 743)  
 Yu, J., Whitelaw, C. A., Nierman, W. C., Bhatnagar, D. and Cleveland, T. E.  
 Aspergillus flavus expressed sequence tags for identification of genes with putative roles in aflatoxin contamination of crops  
 FEMS Microbiol. Lett. (2004) In press  
 CONTACT: Yu J  
 JOURNAL Food and Feed Safety Research Unit  
 USDA/ARS, Southern Regional Research Center  
 1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA  
 Tel: 504 286 4405  
 Fax: 504 286 4419

Email: jinyu@arrrc.ars.usda.gov  
Contact Dr. Yu at USDA/ARS SRRC (jinyu@arrrc.ars.usda.gov) for clone  
information  
PCR Primers  
FORWARD: M13F  
BACKWARD: M13R  
Seq primer: M13 Forward  
POLYA=No.

FEATURES  
source

Location/Qualifiers  
1. 743  
/organism="Aspergillus flavus"  
/mol\_type="mRNA"  
/strain="NRRL 3357"  
/db\_xref="taxon:5059"  
/clone="NAFCU57"  
/sex="asexual mycelia"  
/cell\_type="mycelia"  
/dev\_stage="developmental stages from 18 to 96 hours"  
/lab\_host="E. coli DH10B T1 resistant cells"  
/clone\_1ib="Aspergillus flavus Normalized cDNA Expression  
Library"  
/note="Vector: pBluescript (SK+) (Stratagene), antibiotic  
selection marker: Carbenicillin; Site\_1: NotI, at the 5  
prime end; Site\_2: EcoRI, at the 3 prime end; This  
normalized cDNA expression library was constructed using a  
mixture of mycelial cells grown under eight different  
medium conditions and harvested at 5 time points (18, 24,  
48, 72, 96 hours). The poly-A sequence was trimmed off  
before ligating to vector."

ORIGIN

Query Match 5.3%; Score 44.4; DB 7; Length 743;  
Best Local Similarity 49.2%; Pred. No. 1.1;  
Matches 117; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
QY 579 CCATGTCAGAACGATGCTCCACTCAAGCATCGGTACGAGACAGCGCATGATCT 638  
DB 424 CAAGGCGCGGCCCAAGACACCTCTTAAGATCCAGATGTTGTTCCGAAGCCAAACA 483  
QY 639 GCACACCCCGGGGTGTGTCCTGTGTCGCGGAGGGTAATTCCTCCGCTGGGGTAGC 698  
DB 484 GCTCGCCCCCAACGCGAGTGTGGGTCTGTGTCGATCTGGGCTACTGCTGGGGTGA 543  
QY 699 GCTCACTCCACGCTCGCGGCAAGAGACGCCAGCATCCCACTGCGACAAATACAGCCCA 758  
DB 544 GATCACCGGCTGCGCTCGGGCAAGGACAACAAGATCTTCAAGCGCGCGTGAAGTCCA 603  
QY 759 GGTGATTTGCTGCTTGGGGCGGCTGCTTCTGCTCCGTATGTACGTGGGGATCTC 816  
DB 604 CCCCGTATGCTTGGCCCAAGATGCGAAGTGTGCTCAATCCCATGCGCGTGTCTC 661

RESULT 14

LOCUS CO143444 548 bp mRNA linear EST 17-JUN-2004  
DEFINITION EST38115 Aspergillus flavus Normalized cDNA Expression Library  
ACCESSION CO143444  
VERSION CO143444.1 GI:48897445  
KEYWORDS EST.  
SOURCE Aspergillus flavus  
ORGANISM Aspergillus flavus  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
1 (bases 1 to 548)  
REFERENCE 1 (bases 1 to 548)  
AUTHORS Yu, J., Whitefaw, C.A., Nierman, W.C., Bhatnagar, D. and Cleveland, T.E.  
TITLE Aspergillus flavus expressed sequence tags for identification of  
genes with putative roles in aflatoxin contamination of crops  
JOURNAL FEMS Microbiol. Lett. (2004) In press  
COMMENT Food and Feed Safety Research Unit  
USDA/ARS, Southern Regional Research Center  
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA

Tel: 504 286 4405  
Fax: 504 286 4419  
Email: jinyu@arrrc.ars.usda.gov  
Contact Dr. Yu at USDA/ARS SRRC (jinyu@arrrc.ars.usda.gov) for clone  
information  
PCR Primers  
FORWARD: M13F  
BACKWARD: M13R  
Seq primer: M13 Forward  
POLYA=No.

FEATURES  
source

Location/Qualifiers  
1. 548  
/organism="Aspergillus flavus"  
/mol\_type="mRNA"  
/strain="NRRL 3357"  
/db\_xref="taxon:5059"  
/clone="NAFPB15"  
/sex="asexual mycelia"  
/cell\_type="mycelia"  
/dev\_stage="developmental stages from 18 to 96 hours"  
/lab\_host="E. coli DH10B T1 resistant cells"  
/clone\_1ib="Aspergillus flavus Normalized cDNA Expression  
Library"  
/note="Vector: pBluescript (SK+) (Stratagene), antibiotic  
selection marker: Carbenicillin; Site\_1: NotI, at the 5  
prime end; Site\_2: EcoRI, at the 3 prime end; This  
normalized cDNA expression library was constructed using a  
mixture of mycelial cells grown under eight different  
medium conditions and harvested at 5 time points (18, 24,  
48, 72, 96 hours). The poly-A sequence was trimmed off  
before ligating to vector."

ORIGIN

Query Match 5.2%; Score 43.6; DB 7; Length 548;  
Best Local Similarity 49.1%; Pred. No. 1.7;  
Matches 115; Conservative 0; Mismatches 119; Indels 0; Gaps 0;  
QY 583 GTCAAGAACGATGCTCCACTCAAGCATCGGTACGAGACAGCGCATGATGAC 642  
DB 1 GCCCGGCCCAAGACACCTCTTAAGATCCAGATGTTGTTCCGAAGCCAAAGCTC 60  
QY 643 ACCCGGGGTGTGTCCTGTGTCGCGGAGGTAATTCCTCCGCTGCTGGTAGCGCTC 702  
DB 61 GCCCCCAACGCGAGATTTGGGTCTGTGTCGATCTCGGCTACTGCTGGGGTGA 120  
QY 703 ACTCCACGCTCGCGGCAAGAGACGCAGATCCCACTGCGACAAATACAGCCACGTC 762  
DB 121 ACCGCTCTGGCTCGGCGAAGGACAACAAGATCTTCAAGCGCGCGTCAATGCCACCCC 180  
QY 763 GATTTGCTGTTGGGCGGCTGCTTCTGCTCCGTATGTACGTGGGGATCTC 816  
DB 181 GCTATGCTTGGCCCAAGATGCGAAGTGTGCTCAATCCCATGCGCGTGTCTC 234

RESULT 15

LOCUS CD109759 899 bp mRNA linear EST 15-MAY-2003  
DEFINITION AGENCOURT.13988238 NIH MGC.147 Homo sapiens cDNA clone  
IMAGE:30347093 5', mRNA sequence.  
ACCESSION CD109759  
VERSION CD109759.1 GI:30753968  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 899)  
REFERENCE 1 (bases 1 to 899)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Stefan Hansson

CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>

Plate: NDM387 row: k column: 06  
 High quality sequence stop: 265.

# FEATURES

Location/Qualifiers  
 1..899  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30347093"  
 /issue\_type="Human Placenta"  
 /lab\_host="DH10B TONA"  
 /clone\_lib="NIH\_MGC\_147"  
 /note="Organ: Placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamH; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: This is a NIH\_MGC library."

## ORIGIN

Query Match 5.2%; Score 43.2; DB 6; Length 899;  
 Best Local Similarity 47.1%; Pred. No. 2.4;

Matches 132; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

```

QY 174 CAAGGCTCGCACGCCGAGGCGAGGCGCTGAGCCCGGTACCTTGGCCCTTA 233
    |||||
DB 24  CCGAGGCTGCCCCGCTGGGCCCCCGGAGGAGCGGCGCTACACCTTCGCCCGGCGC 83
    |||||
QY 234 TGGCAACGAGGCGATGGGGTGGGCGAGATGCTCTGTACCCCGTGGCTCCCGGCTAG 293
    |||||
DB 84  TGGGAGGGCGGCGAGTGGGTGCAACCGGACCCCTGCCCCCTGGGCGCCCTCCCG 143
    |||||
QY 294 TTGGGCCCCCAGTACCCCGCGGTAGTGCCTAAATTGGTAAAGTCATCGATACCT 353
    |||||
DB 144 CGGGCGGTGAGTACCCCTGCGAGCGGCTCCCTAGCTAAGAGGAGCGGAGGGGGCTC 203
    |||||
QY 354 CACATCGGCTTGGCGACTCTCATGGGGTACATTCGCTGTCGGGCGCTCCCTTAGGGG 413
    |||||
DB 204 CGGGCGCGCGAGCAGACCTGCTCCGGCGCGGCTCGCGCTGCTCTCCGGAGCGG 263
    |||||
QY 414 CGGTGCGAGGGCCCTGGCGGATGGCGTCCGGGTTCTGGAG 453
    |||||
DB 264 CAGCCCTTAGCCCGGGCGGCGAGGGCTGGGGGTTCTTCGAG 303
    |||||
  
```

Search completed: February 21, 2005, 04:05:44  
 Job time : 3567 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 01:19:35 ; Search time 194 Seconds

(without alignments)  
7034.303 Million cell updates/sec

Title: US-09-664-363-5

Sequence: 1 AAAAAAACCAAGTAACAC.....TTCGCGATCTGTTTCCCG 834

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued\_Patenta\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	834	100.0	834	3 US-08-191-160-5	Sequence 5, Appl1
2	802.6	96.2	2116	3 US-08-191-160-21	Sequence 21, Appl1
3	737	88.4	2433	3 US-08-612-973-49	Sequence 49, Appl1
4	737	88.4	2433	3 US-08-927-597-49	Sequence 49, Appl1
5	733.8	88.0	9413	4 US-09-827-688-6	Sequence 6, Appl1
6	730.6	87.6	1539	2 US-08-470-426B-17	Sequence 17, Appl1
7	730.6	87.6	1863	2 US-08-470-426B-14	Sequence 14, Appl1
8	730.6	87.6	9595	3 US-09-014-416-4	Sequence 4, Appl1
9	730.6	87.6	9599	3 US-09-014-416-6	Sequence 6, Appl1
10	730.6	87.6	11076	4 US-09-539-601-1	Sequence 1, Appl1
11	730.6	87.6	11076	4 US-09-539-601-19	Sequence 19, Appl1
12	730.6	87.6	11076	4 US-09-539-601-25	Sequence 25, Appl1
13	730.6	87.6	11076	4 US-09-539-601-31	Sequence 31, Appl1
14	719.4	86.3	1037	1 US-08-462-195-1	Sequence 1, Appl1
15	719.4	86.3	1037	2 US-08-636-883-1	Sequence 1, Appl1
16	719.4	86.3	1037	3 US-09-127-829-1	Sequence 1, Appl1
17	716.2	85.9	1167	1 US-08-324-977-9	Sequence 9, Appl1
18	716.2	85.9	1167	2 US-08-384-616-9	Sequence 9, Appl1
19	716.2	85.9	1167	2 US-08-904-686A-9	Sequence 9, Appl1
20	716.2	85.9	1167	3 US-09-315-850-3	Sequence 9, Appl1
21	716.2	85.9	1499	1 US-08-324-977-3	Sequence 3, Appl1
22	716.2	85.9	1499	1 US-08-324-977-3	Sequence 3, Appl1
23	716.2	85.9	1499	2 US-08-904-686A-3	Sequence 3, Appl1
24	716.2	85.9	1499	3 US-09-315-850-3	Sequence 3, Appl1
25	716.2	85.9	6039	1 US-08-324-977-11	Sequence 11, Appl1
26	716.2	85.9	6039	2 US-08-384-616-11	Sequence 11, Appl1
27	716.2	85.9	6039	2 US-08-904-686A-11	Sequence 11, Appl1

28	716.2	85.9	6039	3 US-09-315-850-11	Sequence 11, Appl1
29	716.2	85.9	9030	1 US-08-324-977-13	Sequence 13, Appl1
30	716.2	85.9	9030	2 US-08-384-616-13	Sequence 13, Appl1
31	716.2	85.9	9030	2 US-08-904-686A-13	Sequence 13, Appl1
32	716.2	85.9	9030	3 US-09-315-850-13	Sequence 13, Appl1
33	716.2	85.9	9416	1 US-08-324-977-1	Sequence 1, Appl1
34	716.2	85.9	9416	2 US-08-384-616-1	Sequence 1, Appl1
35	716.2	85.9	9416	2 US-08-904-686A-1	Sequence 1, Appl1
36	716.2	85.9	9416	3 US-09-315-850-1	Sequence 1, Appl1
37	716.2	85.9	9416	3 US-08-823-895A-27	Sequence 27, Appl1
38	713	85.5	1037	1 US-08-462-195-3	Sequence 3, Appl1
39	713	85.5	1037	2 US-08-636-883-3	Sequence 3, Appl1
40	713	85.5	1037	3 US-09-127-829-3	Sequence 3, Appl1
41	711.4	85.3	9472	4 US-08-150-204E-96	Sequence 96, Appl1
42	594.6	71.3	12980	3 US-08-811-566-5	Sequence 5, Appl1
43	594.6	71.3	12980	3 US-09-034-756-5	Sequence 5, Appl1
44	593	71.1	9365	4 US-09-827-688-7	Sequence 7, Appl1
45	593	71.1	9401	2 US-08-432-693-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-191-160-5  
; Sequence 5, Application US/08191160  
; Patent No. 6210675  
; GENERAL INFORMATION:  
; APPLICANT: Highfield, Peter Edmund  
; APPLICANT: Rodgers, Brian Collin  
; APPLICANT: Tedder, Richard Seton  
; APPLICANT: Barbara, John Anthony James  
; TITLE OF INVENTION: Viral Agent  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Piggy, Ernst & Kurz  
; STREET: 1700 K Street  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage  
; COMPUTER: IBM AT compatible  
; OPERATING SYSTEM: MS-DOS V3.2  
; SOFTWARE: wordperfect 5.0 (DOS text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/191,160  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/628,516  
; FILING DATE: 17 DEC 1990  
; APPLICATION NUMBER: UK 89 28 562.1  
; FILING DATE: 18 DEC 1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 414.0  
; FILING DATE: 27 FEB 1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 814.1  
; FILING DATE: 03 MAR 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: E. Anthony F199  
; REGISTRATION NUMBER: 27,195  
; REFERENCE/DOCKET NUMBER: 1645-103A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 833-5740  
; TELEFAX: (202) 833-5744  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 834 base pairs  
; TYPE: nucleotide with corresponding protein  
; STRANDEDNESS: single

```

/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA to genomic RNA
/ ORIGINAL SOURCE:
/ ORGANISM: human; serum infectious for PT-NANBH
/ IMMEDIATE SOURCE:
/ LIBRARY: clone BR11 from cDNA library in lambda gt11
/ FEATURE:
/ LOCATION: from 1 to 834 bp portion of the PT-NANBH
/ LOCATION: polyprotein
/ OTHER INFORMATION: probably encodes viral structural
/ OTHER INFORMATION: proteins
US-08-191-160-5

```

```

Query Match          100.0%; Score 834; DB 3; Length 834;
Best Local Similarity 100.0%; Pred. No. 7.5e-245;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AGAAAAACCAACGTAACCAACCTCCGCCACAGACGTGAGTTCCGGGCGGTGAT 60
DB 1 AGAAAAACCAACGTAACCAACCTCCGCCACAGACGTGAGTTCCGGGCGGTGAT 60
QY 61 CAGATCGTTGGTGAAGTTAACTGTTGCGGCGAGGGGCGCCAGGTTGGGTGGCGCG 120
DB 61 CAGATCGTTGGTGAAGTTAACTGTTGCGGCGAGGGGCGCCAGGTTGGGTGGCGCG 120
QY 121 ACTAGGAAGACTTCGAGCGCGTGCACACCTCGTGAAGGCGCAACTATCCCAAGGCT 180
DB 121 ACTAGGAAGACTTCGAGCGCGTGCACACCTCGTGAAGGCGCAACTATCCCAAGGCT 180
QY 181 CGCAGCGCCGAGGCGAGGCGCTGGGCTCAAGCCGGGATACCTTGGCCCTTATGGCAGC 240
DB 181 CGCAGCGCCGAGGCGAGGCGCTGGGCTCAAGCCGGGATACCTTGGCCCTTATGGCAGC 240
QY 241 GAGGCGATGGGGGCGAGGATGGCTCCGTGACCCCGTGGGCTCCCGGCTGATTGGGGCG 300
DB 241 GAGGCGATGGGGGCGAGGATGGCTCCGTGACCCCGTGGGCTCCCGGCTGATTGGGGCG 300
QY 301 CCCACTGACCCCGGCGTGAAGTTCGAGTAATTTGGTAAGATCATGATACCTTCACATGC 360
DB 301 CCCACTGACCCCGGCGTGAAGTTCGAGTAATTTGGTAAGATCATGATACCTTCACATGC 360
QY 361 GCGTTGCGCGACTTCATAGGGGTAATTCGCTCGCGCGCTCCCTTAAGGGGCGCTGCC 420
DB 361 GCGTTGCGCGACTTCATAGGGGTAATTCGCTCGCGCGCTCCCTTAAGGGGCGCTGCC 420
QY 421 AAGGCGCTGGCGATGGCGTCCGGGTTCTGAGAGAGGCGGTAACTATCAACAGGGAAT 480
DB 421 AAGGCGCTGGCGATGGCGTCCGGGTTCTGAGAGAGGCGGTAACTATCAACAGGGAAT 480
QY 481 TTACCCGGTGGCTTTCTCTATCTTCTCTTGGCTTTCGTGCTGTTGACATTCGA 540
DB 481 TTACCCGGTGGCTTTCTCTATCTTCTCTTGGCTTTCGTGCTGTTGACATTCGA 540
QY 541 GCTTCGCGTTATGAGAGTGGCGAAGCTGTCCGGGATTTACATGTCAAGCAATTCCTCC 600
DB 541 GCTTCGCGTTATGAGAGTGGCGAAGCTGTCCGGGATTTACATGTCAAGCAATTCCTCC 600
QY 601 AACTCAAGATGTGTAGAGACAGCGGACATGATATCAACCCCGGGGTGTGGCC 660
DB 601 AACTCAAGATGTGTAGAGACAGCGGACATGATATCAACCCCGGGGTGTGGCC 660
QY 661 TGTGTCCGGAGGTAATTCCTCCGCTGTGGGTAGCGCTCACTCCACGCTCGCGGCG 720
DB 661 TGTGTCCGGAGGTAATTCCTCCGCTGTGGGTAGCGCTCACTCCACGCTCGCGGCG 720
QY 721 AAGAGCGCAGATCCCACTGCGAACAATACGACGCGCATGATTCCTGTTGGGGCG 780
DB 721 AAGAGCGCAGATCCCACTGCGAACAATACGACGCGCATGATTCCTGTTGGGGCG 780
QY 781 GGTGCTTCTGCTGCGCTATGTAAGTGGGGGATCTCTGGGATTCGTTTCCCG 834
DB 781 GGTGCTTCTGCTGCGCTATGTAAGTGGGGGATCTCTGGGATTCGTTTCCCG 834

```

# RESULT 2 US-08-191-160-21

```

/ Sequence 21, Application US/08191160
/ Patent No. 6210675
/ GENERAL INFORMATION:
/ APPLICANT: Highfield, Peter Edmund
/ APPLICANT: Rodgers, Brian Colin
/ APPLICANT: Tedder, Richard Seton
/ APPLICANT: Barbara, John Anthony James
/ TITLE OF INVENTION: Viral Agent
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Rothwell, Figg, Ernst & Kurz
/ STREET: 1700 K Street
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20006
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
/ COMPUTER: IBM AT compatible
/ OPERATING SYSTEM: MS-DOS V3.2
/ SOFTWARE: Wordperfect 5.0 (DOS text)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/191,160
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/628,516
/ FILING DATE: 17 DEC 1990
/ APPLICATION NUMBER: UK 89 28 562.1
/ FILING DATE: 18 DEC 1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: UK 90 04 414.0
/ FILING DATE: 27 FEB 1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: UK 90 04 814.1
/ FILING DATE: 03 MAR 1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: E. Anthony Figg
/ REGISTRATION NUMBER: 27,195
/ REFERENCE/DOCKET NUMBER: 1645-103A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 833-5740
/ TELEFAX: (202) 833-5744
/ INFORMATION FOR SEQ ID NO: 21:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2116 base pairs
/ TYPE: nucleotide with corresponding protein
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA to genomic RNA
/ ORIGINAL SOURCE:
/ ORGANISM: human; serum infectious for PT-NANBH
/ IMMEDIATE SOURCE:
/ LIBRARY: cDNA clones from 5' end of the genome
/ FEATURE:
/ LOCATION: from 308 to 2116 bp start of the PT-NANBH
/ LOCATION: polyprotein
/ OTHER INFORMATION: viral structural and non-structural
/ OTHER INFORMATION: proteins
US-08-191-160-21

```

```

Query Match          96.2%; Score 802.6; DB 3; Length 2116;
Best Local Similarity 99.3%; Pred. No. 4.9e-235;
Matches 827; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

```

QY 1 AGAAAAACCAACGTAACCAACCTCCGCCACAGACGTGAGTTCCGGGCGGTGAT 60
DB 332 AGAAAAACCAACGTAACCAACCTCCGCCACAGACGTGAGTTCCGGGCGGTGAT 391
QY 61 CAGATCGTTGGTGAAGTTAACTGTTGCGGCGAGGGGCGCCAGGTTGGGTGGCGCG 120

```

|||||  
Db 392 CAGATGTTGTTGAGTTTACCTGTTGCGCGCAGAGGGGCCCCAGGTTGGGTGTGGCGGG 451  
Qy 121 ACTAGGAAGACTTCCAGGCGGTGCAACTCTGTGAAAGGCGACACTTATCCCAAGGCT 180  
Db 452 ACTAGGAAGACTTCCAGGCGGTGCAACTCTGTGAAAGGCGACACTTATCCCAAGGCT 511  
Qy 181 CGCAGACCCGAGGCGAGGGGCTGGGCTCAGCCCGGGTACCTTGGGCCCCCTATATGCAAC 240  
Db 512 CGCAGACCCGAGGCGAGGGGCTGGGCTCAGCCCGGGTACCTTGGGCCCCCTATATGCAAC 571  
Qy 241 GAGGGCATGGGTGGGCGAGATGGCTCTGTCAACCCCGTGGCTCCCGGCTATGTTGGGCG 300  
Db 572 GAGGGCATGGGTGGGCGAGATGGCTCTGTCAACCCCGTGGCTCCCGGCTATGTTGGGCG 631  
Qy 301 CCCACTGACCCCGGCGGTAGTGGCTATTTGGGTAAATCATCATGATCCTCATATGC 360  
Db 632 CCCACTGACCCCGGCGGTAGTGGCTATTTGGGTAAATCATCATGATCCTCATATGC 691  
Qy 361 GGGTGGCGGACTCTGATGGGGTCAATTCGCGCTGTCGCGGCTCCCTTA-GGGGCGCTGC 419  
Db 692 GGGTGGCGGACTCTGATGGGGTCAATTCGCGCTGTCGCGGCTCCCTTA-GGGGCGCTGC 750  
Qy 420 CAGGGCCCTGGCGCATGGGCTCCGGGTTCTGAGAGACGGCGTGAATGCAACAGGGAA 479  
Db 751 CAGGGCCCTGGCGCATGGGCTCCGGGTTCTGAGAGACGGCGTGAATGCAACAGGGAA 810  
Qy 480 TTATCCCGGTGCTCTTTCTCTATCTTCTCTTGGCTTGGCTGCTCTGTTTGAACATTC 539  
Db 811 TTATCCCGGTGCTCTTTCTCTATCTTCTCTTGGCTTGGCTGCTCTGTTTGAACATTC 870  
Qy 540 AGCTTCGGCTTAAAGTGGCGGCAAGTTCGCGGAACTTACATGTAACGAACGATGGCTC 599  
Db 871 AGCTTCGGCTTAAAGTGGCGGCAAGTTCGCGGAACTTACATGTAACGAACGATGGCTC 930  
Qy 600 CAACCTCAAGCATGTGTAGAGAGACAGGAGCATGATCATGACACCCCGGGGTGTGGC 659  
Db 931 CAACCTCAAGCATGTGTAGAGAGACAGGAGCATGATCATGACACCCCGGGGTGTGGC 990  
Qy 660 CTGTGTCCGAGGAGTAATTCCTCCGCTGCTGGGTAGGCGCTCACTCCACGCTGCGGC 719  
Db 991 CTGTGTCCGAGGAGTAATTCCTCCGCTGCTGGGTAGGCGCTCACTCCACGCTGCGGC 1050  
Qy 720 CAGAGACCGCAAGATATCCCACTGCGCAATATGACAGCCCAAGTGGATTTGCTGTGGGCG 779  
Db 1051 CAGAGACCGCAAGATATCCCACTGCGCAATATGACAGCCCAAGTGGATTTGCTGTGGGCG 1110  
Qy 780 GGGTGGCTTCTGCTCGCTATGATGATGGGGGATCTCTGGGATCTGTTTTC 832  
Db 1111 GGGTGGCTTCTGCTCGCTATGATGATGGGGGATCTCTGGGATCTGTTTTC 1163

RESULT 3  
US-08-612-973-49  
; Sequence 49. Application us/08612973  
; Patent No. 6150134  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEBERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARTYNOFF, GUY  
; APPLICANT: BUYSE, MARIE-ANGY  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P. C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
FILING DATE: 11-MAR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BYRNE, THOMAS E.  
REGISTRATION NUMBER: 32,205  
REFERENCE/DOCKET NUMBER: 1487-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ. ID NO.: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2433 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2430  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..2427  
US-08-612-973-49

Query Match 88.4%; Score 737; DB 3; Length 2433;

Best Local Similarity 94.4%; Pred. No. 6,2e-215; Matches 786; Conservative 0; Mismatches 45; Indels 2; Gaps 2;

Qy 1 AGAAAAACAAAGTAACACCAACTCCGCCACAGAGTGAAGTTCGCGGCGGTGT 60  
Db 25 AGAAAAACAAAGTAACACCAACTCCGCCACAGAGTGAAGTTCGCGGCGGTGT 84  
Qy 61 CAGATGTTGTTGAGTAACTCTGTTGCGCGGAGGGGCCCCAGTGGGTGTGCGCGG 120  
Db 85 CAGATGTTGTTGAGTAACTCTGTTGCGCGGAGGGGCCCCAGTGGGTGTGCGCGG 144  
Qy 121 ACTAGGAAGACTTCCAGGCGGTGCAACTCTGTGAAAGGCGCAACTATCCCAAGGCT 180  
Db 145 ACTAGGAAGACTTCCAGGCGGTGCAACTCTGTGAAAGGCGCAACTATCCCAAGGCT 204  
Qy 181 CGCAGACCCGAGGCGAGGGGCTGGGCTCAGCCCGGGTACCTTGGGCCCCCTATATGCAAC 240  
Db 205 CGCAGACCCGAGGCGAGGGGCTGGGCTCAGCCCGGGTACCTTGGGCCCCCTATATGCAAC 264  
Qy 241 GAGGGCATGGGTGGGCGAGATGGCTCTGTCAACCCCGTGGCTCCCGGCTATGTTGGGCG 300  
Db 265 GAGGGCATGGGTGGGCGAGATGGCTCTGTCAACCCCGTGGCTCCCGGCTATGTTGGGCG 324  
Qy 301 CCCACTGACCCCGGCGGTAGTGGCTATTTGGGTAAATCATCATGATCCTCATATGC 360  
Db 325 CCCACTGACCCCGGCGGTAGTGGCTATTTGGGTAAATCATCATGATCCTCATATGC 384  
Qy 361 GGGTGGCGGACTCTGATGGGGTCAATTCGCGCTGTCGCGGCTCCCTTA-GGGGCGCTGC 419  
Db 385 GGGTGGCGGACTCTGATGGGGTCAATTCGCGCTGTCGCGGCTCCCTTA-GGGGCGCTGC 443  
Qy 420 CAGGGCCCTGGCGCATGGGCTCCGGGTTCTGAGAGACGGCGTGAATGCAACAGGGAA 479  
Db 444 CAGGGCCCTGGCGCATGGGCTCCGGGTTCTGAGAGACGGCGTGAATGCAACAGGGAA 503  
Qy 480 TTATCCCGGTGCTCTTTCTCTATCTTCTCTTGGCTTGGCTGCTCTGTTTGAACATTC 539  
Db 504 TTATCCCGGTGCTCTTTCTCTATCTTCTCTTGGCTTGGCTGCTCTGTTTGAACATTC 563  
Qy 540 AGCTTCGGCTTAAAGTGGCGGCAAGTTCGCGGAACTTACATGTAACGAACGATGGCTC 599

Db	56	AGCTTCGGCTTATGAAAGTGGCAACGTGTCGGGAGATGTCCATGTCACAGAAACGACTGTC	622
Qy	600	CAACTCAAGATCTGTGTACGACACAGCGGACATGATCATGACACACCCCCGGGTGTGTGCC	655
Db	624	CAACTCAAGATGTGTATGATAGGACACGGACATGATCATGACACACCCCCGGGTGTGTGCC	683
Qy	660	CTGTCGTCCGGAGAGGTAAATTCCTCCGGCTGTGGGATAGGCGTCATCTCCACGCTCGCGGC	719
Db	684	CTGCGTTCCGGAGAAACAATCTTCCCGCTGTGGGTAGGGCTCACCCCCACGCTTCGACGC	745
Qy	720	CAGAGACGCAGCATCCCACTGCGCAATACGACGCGCACGTGATTTGCTGTTGGGCG	779
Db	744	TAGGAACGCCAGCGTCCCCACCAACGACATACGACGCGCACGTGATTTGCTGTTGGGCG	803
Qy	780	GGCGCGCTTCGTGTCGCTATGTAGCTGGGGGATCTCGCGGATCTGTTTTCC	832
Db	804	GGCGCGCTTCTGTTCGCTATGTAGCTGGGGGACCTTCGCGGATCTGTTTTCC	856

```

US-08-927-597--49
: Sequence 49, Application US/08927597
: Patent No. 6245503
: GENERAL INFORMATION:
: APPLICANT: MARTENS, GEERT
: APPLICANT: BOSMAN, PONS
: APPLICANT: DE MARTYNOFF, GUY
: APPLICANT: BUYSE, MARIE-ANSE
: TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
: TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
: NUMBER OF SEQUENCES: 111
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHYE P.C.
: STREET: 1100 NORTH GLEBE ROAD
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: U.S.A.
: ZIP: 22201-4714
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EFO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/927,597
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/612,973
: FILING DATE: 11-MAR-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: BYRNE, THOMAS E.
: REGISTRATION NUMBER: 32,205
: REFERENCE/DOCKET NUMBER: 1487-10
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 816-4000
: TELEFAX: (703) 816-4100
: INFORMATION FOR SEQ ID NO: 49:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2433 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..2430
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 1..2427
US-08-927-597-49

```

Query Match	88.4%;	Score 737;	DB 3;	Length 2433;
Best Local Similarity	94.4%;	Pred. No. 6.2e-215;		
Matches 786;	Conservative	0;	Mismatches 45;	Indels 2;
			Gaps	2

QY 1 AGAAAAACAAACGTAAACAAACCACTCCGCCACAGGACGTCAGTTCCCGGCGGTGT 60  
 Db 25 AGAAAAACAAACGTAAACCAACCCGCCGCCACAGGACGTCAAGTTCCCGGCGGTGT 84  
 QY 61 CAGATCGTTGTGTGAAGTTTACCTGTTCCGCCGACAGGGGCCCCAGGTTGGTGTGCCGCG 120  
 Db 85 CAGATCGTTGTGTGAAGTTTACCTGTTCCGCCGACAGGGGCCCCAGGTTGGTGTGCCGCG 144  
 QY 121 ACTAGAAAGACTTCCGAGCGGTGCGCAACTCGTGTGAAGGCGACAACTATCCCCAAGGCT 180  
 Db 145 ACTAGAAAGACTTCCGAGCGGTGCGCAACTCGTGTGAAGGCGACAACTATCCCCAAGGCT 204  
 QY 181 CGCCAGCCCGAAGGCGAGGCTTGGGCTCAGCCCGGGTACCTTTGGCCCTCTATTTGGCAAC 240  
 Db 205 CGCGAGCCCGAAGGCTTGGGCTCAGCCCGGGTACCTTTGGCCCTCTATTTGGCAAT 264  
 QY 241 GAGGGCATTTGGGGTGGGAGGATGGCTCTGTACCCCGTGGCTCCCGGACCTAGTTGGGGG 300  
 Db 265 GAGGGCATTTGGGGTGGGAGGATGGCTCTGTACCCCGGCTCTGGGCTTATTTGGGGG 324  
 QY 301 CCCACTGACCCCGCGCGTAGGTGCGGTAATTTGGGTAAAGTCAATGCAATCCCTCACATGC 360  
 Db 325 CCTACAGACCCCGCGCGTAGGTGCGGTAATTTGGGTAAAGTCAATGCAATCCCTCACATGC 384  
 QY 361 GGCTTGGCGGACCTCTCATTTGGGGTACATTCGCGCTCGTGGGCGTCCCTTA-GGGGGCGGTG 419  
 Db 385 GGCTTGGCGGAC-CTCGTGGGTACATTCGCGCTCGTGGGCGGCCCCCTTAAAGGGCGGTG 443  
 QY 420 CAGGGCCCTGGCGATGGCGTCCGGGTTCTGAGAGACGGCGTGAATGCAACAGGAA 479  
 Db 444 CAGGGCCCTGGCGATGGCGTCCGGGTTCTGAGAGACGGCGTGAATGCAACAGGAA 503  
 QY 480 TTTTACCGGTTGCTCTTCTCTCTATCTTCTCTTGGCTTGGCTGCTGCTGTGACATTTCC 539  
 Db 504 TTTGCCCGGTTGCTCTTCTCTCTATCTTCTCTTGGCTTGGCTGCTGCTGTGACCGTTCC 563  
 QY 540 AGCTTCGCGTTATGAATGGCGCAACGTTCCGGGATCTACATGTGACGAACGATTGCTC 599  
 Db 564 AGCTTCGCGTTATGAATGGCGCAACGTTCCGGGATGTACATGTGACGAACGATGCTC 623  
 QY 600 CAACTCAAGCATGTTGATGAGACAGCGGACATGATTCAGACACCCCGGTTGTGCC 659  
 Db 624 CAACTCAAGCATGTTGATGAGACAGCGGACATGATTCAGACACCCCGGTTGTGCC 683  
 QY 660 CTGTGTCCGGGAGGGTAACTCTCCGCGTGTGGGTAGGCGCTCACTCCACGCTCGCGG 719  
 Db 684 CTGTGTCCGGGAGGAACAACCTCTTCCGCTGTGGGTAGGCGCTCACTCCACGCTCGGAGC 743  
 QY 720 CAAAGACGCCAGCATCCCACTGCGACAAATACAGACGCCAGTGCATTTGCTGTGGGG 779  
 Db 744 TAGGAAGCCAGCGGTCCCACTGCGACAAATACAGACGCCAGTGCATTTGCTGTGGGG 803  
 QY 780 GGCTGCTTCTCGTCCGCTATGTACGTTGGGGGATCTCTGGGATCTGTTTCC 832  
 Db 804 GGCTGCTTCTCGTCCGCTATGTACGTTGGGGGATCTCTGGGATCTGTTTCC 856  
 RESULT 5  
 US-09-827-688--6  
 ; Sequence 6, Application US/09827688  
 ; Patent No. 6821955  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ORSON, FRANK  
 ; APPLICANT: KINSEY, BERNA  
 ; APPLICANT: BHOGAL, BALRA  
 ; TITLE OF INVENTION: MACRODAGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION E  
 ; FILE REFERENCE: P01949US1/1004014



CURRENT APPLICATION NUMBER: US/09/827,688  
CURRENT FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: 60/195,680  
PRIOR FILING DATE: 2000-04-07  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 6  
LENGTH: 9413  
TYPE: DNA  
ORGANISM: HEPATITIS C  
US-09-827-688-6

Query Match 88.0%; Score 733.8; DB 4; Length 9413;  
Best Local Similarity 94.1%; Pred. No. 1.2e-213;  
Matches 784; Conservative 0; Mismatches 47; Indels 2; Gaps 2;

1 AGAAAAACAAAGTAAACCAACCTCCGCCACAGAGCTCAGGTTCCCGGCGGTGGT 60  
354 AGAAAAACAAAGTAAACCAACCTCCGCCACAGAGCTTAAAGTTCCCGGCGGTGGT 413  
61 CAGATGTTGGTGAATTTACCTGTTCCGCCAGAGGGGCCCAAGTTGGTGGCGCGG 120  
414 CAGATGTTGGTGAATTTACCTGTTCCGCCAGAGGGGCCCAAGTTGGTGGCGCGG 473  
121 ACTAGAAAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGCAACCTATCCCAAGGCT 180  
474 ACTAGAAAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGCAACCTATCCCAAGGCT 533  
181 CCGCAGACCCGAGGAGGAGGCGCTGGGCTCAGCCCGGGTACCCTTGGCCCTTATGGCAAC 240  
534 CCGCAGACCCGAGGAGGAGGAGGCGCTGGGCTCAGCCCGGGTACCCTTGGCCCTTATGGCAAC 593  
241 GAGGGGATGGGGGGGAGAGTGGCTCCTGTCAACCCCGGGTCCCGGCGCTACTGGGGC 300  
594 GAGGGGATGGGGGGGAGAGTGGCTCCTGTCAACCCCGGGTCCCGGCGCTACTGGGGC 653  
301 CCCACTGACCCCGGAGGAGTGGCTGCGTAAATTTGGTAAAGTACGATACCTCAGATGC 360  
654 CCCACAGACCCCGGAGGAGTGGCTGCGTAAATTTGGTAAAGTACGATACCTCAGATGC 713  
361 GCGCTTGGCGGACTCTCATGAGGAGTAACTTCGCTGGCGGCTCCCTTA-GGGGCGCTGC 419  
714 GCGCTTGGCGGACTCTCATGAGGAGTAACTTCGCTGGCGGCTCCCTTA-GGGGCGCTGC 772  
420 CAGGGGCTTGGCGGAGTGGCTCCGGGTTCTGAGAGAGCGGTGAATTAAGCAAGAGAA 479  
773 CAGGGGCTTGGCGGAGTGGCTCCGGGTTCTGAGAGAGCGGTGAATTAAGCAAGAGAA 832  
480 TTAAACCGGTTGCTCTTCTCTATCTTCTTGGCTTGGCTTGGTGAACAATGC 539  
833 TCTGCCCGGTTGCTCTTCTCTATCTTCTTGGCTTGGCTTGGTGAACAATGC 892  
540 AGCTTCGCTTATGAAGTGGCAACGTTGTCGGGATCTACATGTACAGAACGATTTGCTC 599  
893 AGCTTCGCTTATGAAGTGGCAACGTTGTCGGGATCTACATGTACAGAACGATTTGCTC 952  
600 CAATCTAACATCTGTATGAGAGACAGCGAATGATACGACACCCCGGGGTGCTGCTC 659  
953 CAATCTAACATCTGTATGAGAGACAGCGAATGATACGACACCCCGGGGTGCTGCTC 1012  
660 CTGTGTCGGGAGGAGTAAATTTCTCCGCTGCTGGTGGGCTCACTCCCAAGCTCGCGG 719  
1013 CTGTGTCGGGAGGAGTAAATTTCTCCGCTGCTGGTGGGCTCACTCCCAAGCTCGCGG 1072  
720 CAAGAGCGCAGAGTCCCACTGCGCAATACGACGCAACGTCGATTTGCTGTTGGGGC 779  
1073 CAAGAGCGCAGAGTCCCACTGCGCAATACGACGCAACGTCGATTTGCTGTTGGGGC 1132  
780 GCGTGGCTTCTGCTCCGCTATGATGAGTGGGAGATCTCTGGGATCTGTTTCC 832  
1133 GCGTGGCTTCTGCTCCGCTATGATGAGTGGGAGATCTCTGGGATCTGTTTTC 1185

RESULT 6  
US-08-470-426B-17  
Sequence 17, Application US/08470426B  
Patent No. 5856458  
GENERAL INFORMATION:  
APPLICANT: Okamoto, Hiroaki  
TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR  
TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B  
TITLE OF INVENTION: HEPATITIS VIRUS  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Beveridge, Degrandi, Wellacher & Young,  
ADDRESSEE: L.L.P.  
STREET: 1850 M Street, N.W., Suite 800  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,426B  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-153402  
FILING DATE: 12-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Wellacher, Robert G.  
REGISTRATION NUMBER: 20,531  
REFERENCE/DOCKET NUMBER: 06/59-47083.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2811  
TELEFAX: (202) 659-1462  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1539 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
US-08-470-426B-17.

Query Match 87.6%; Score 730.6; DB 2; Length 1539;  
Best Local Similarity 93.9%; Pred. No. 4.4e-213;  
Matches 782; Conservative 0; Mismatches 49; Indels 2; Gaps 2;

1 AGAAAAACAAAGTAAACCAACCTCCGCCACAGAGCTCAGGTTCCCGGCGGTGGT 60  
25 AGAAAAACAAAGTAAACCAACCTCCGCCACAGAGCTTAAAGTTCCCGGCGGTGGT 84  
61 CAGATGTTGGTGAATTTACCTGTTCCGCCAGAGGGGCCAAGTTGGTGGCGCGG 120  
85 CAGATGTTGGTGAATTTACCTGTTCCGCCAGAGGGGCCAAGTTGGTGGCGCGG 144  
121 ACTAGAAAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGCAACCTATCCCAAGGCT 180  
145 ACTAGAAAGACTTCCGAGCGGTGCAACCTCGTGAAGGCGCAACCTATCCCAAGGCT 204  
205 CCGCAGACCCGAGGAGGAGGCGCTGGGCTCAGCCCGGGTACCCTTGGCCCTTATGGCAAT 264  
241 GAGGGGATGGGGGGGAGAGTGGCTCCTGTGAACCCCGGGTCCCGGCTATGGGGG 300  
265 GAGGGGATGGGGGGGAGAGTGGCTCCTGTGAACCCCGGGTCCCGGCTATGGGGG 324  
301 CCCACTGACCCCGGAGTGGTGGCTAAATTTGGTAAAGTACGATACCTCAGATGC 360



```
Patent No. 6153421
GENERAL INFORMATION:
APPLICANT: Yanagi, Masayuki
APPLICANT: Buhk, Jens
APPLICANT: Emerson, Susanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
FILE REFERENCE: 20264276
CURRENT APPLICATION NUMBER: US/09/014,416
CURRENT FILING DATE: 1998-01-27
EARLIER APPLICATION NUMBER: US 60/053,062
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 65
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 9595
TYPE: DNA
ORGANISM: Hepatitis C virus
US-09-014-416-4
```

```
Query Match      87.6%; Score 730.6; DB 3; Length 9595;
Best Local Similarity 93.9%; Pred. No. 1.1e-212;
Matches 782; Conservative 0; Mismatches 49; Indels 2; Gaps 2;
```

```
QY 1 AAAAAAAAAAAGTAACCAACCTCCGCGCAGAGAGTGTCCCGGGCGGTGT 60
DB 366 AAAAAAAAAAAGTAACCAACCTCCGCGCAGAGAGTGTCCCGGGCGGTGT 425
QY 61 CAGATCGTGTGTGAGTTTAACTGTGTGCGCGCAGAGGCGCCAGTTGGGTGTGCGCGG 120
DB 426 CAGATCGTGTGTGAGTTTAACTGTGTGCGCGCAGAGGCGCCAGTTGGGTGTGCGCGG 485
QY 121 ACTAGGAAGACTTCCGAGCGGTGTGCAACTGTGTGAAGAGCGACAACCTATCCCAAGGCT 180
DB 486 ACTAGGAAGACTTCCGAGCGGTGTGCAACTGTGTGAAGAGCGACAACCTATCCCAAGGCT 545
QY 181 CCGCAGCGCGAGGCGAGGCGCTGCGGCTCAGCCCGGGTACCTTGGCCCTTATGCGCAGC 240
DB 546 CCGCAGCGCGAGGCGAGGCGCTGCGGCTCAGCCCGGGTACCTTGGCCCTTATGCGCAGC 605
QY 241 GAGGCGATGGGCGAGGCGATGGCTCTGTCAACCCCGTGGCTCCCGGCTATGGGGGCT 300
DB 606 GAGGCGATGGGCGAGGCGATGGCTCTGTCAACCCCGTGGCTCCCGGCTATGGGGGCT 665
QY 301 CCGACTGACCCCGGCGGTAGTGGGTAAATTGGGTAAAGTCAATGATCCTTACATGTC 360
DB 666 CCGACTGACCCCGGCGGTAGTGGGTAAATTGGGTAAAGTCAATGATCCTTACATGTC 725
QY 361 GCGTTCGCGCATCTTCATAGGGGTACATTCGCTGCTGCGGCTCCCTTA-GGGGGCGCTGC 419
DB 726 GCGTTCGCGCAT-TCATATGGGTACATTCGCTGCTGCGGCTCCCTTAAGGGGGCGCTGC 784
QY 420 CAGGCGCTTGCGCATGGCTCGGCGTTCGAGAGACGCGGTGAATATATGCAACAGGAA 479
DB 785 CAGGCGCTTGCGCATGGCTCGGCGTTCGAGAGACGCGGTGAATATATGCAACAGGAA 844
QY 480 TTATACCGGTTGCTCTTTCTATCTTCTCTTGGCTTTGCTGCTGTTTGAACATTC 539
DB 845 CTTCGCGGTTGCTCTTTCTATCTTCTCTTGGCTTTGCTGCTGTTTGAACATTC 904
QY 540 AGCTTCGCGTTATGAGTGGGAGAAAGTTCGCGGATCTACATGTCACGAACATGTC 599
DB 905 AGCTTCGCGTTATGAGTGGGAGAAAGTTCGCGGATCTACATGTCACGAACATGTC 964
QY 600 CAACCTCAACATGTTAGAGAGACAGCGGATGATCATGACACCCCGGCGGTGTGC 659
DB 965 CAACCTCAACATGTTAGAGAGAGAGCGGATGATCATGACACCCCGGCGGTGTGC 1024
QY 660 CTGTGTCCGAGAGGTAATTCCTCCGCTGCTGGGTAGCGCTCACTCCAGCTCGCGGC 719
DB 1025 CTGTGTCCGAGAGGTAATTCCTCCGCTGCTGGGTAGCGCTCACTCCAGCTCGCGGC 1084
```

```
QY 720 CAGAGCGCAGCATCCCACTGCGACATATGACAGCGCAGTGTGATTTGCTGTGGGCG 779
DB 1085 CAGAGATGCGACGCTCCCACTACGACATATGAGCGCAGTGTGATTTGCTGTGGGAGC 1144
QY 780 GCGTGCCTTCTGCTCGCTATGATGAGTGGGGATCTTCGCGATCTGTTTCC 832
DB 1145 GCGTGCCTTCTGCTCGCTATGATGAGTGGGGATCTTCGCGATCTATTTTCC 1197
```

## RESULT 9

```
US-09-014-416-6
Sequence 6, Application US/09014416
Patent No. 6153421
GENERAL INFORMATION:
APPLICANT: Yanagi, Masayuki
APPLICANT: Buhk, Jens
APPLICANT: Emerson, Susanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
FILE REFERENCE: 20264276
CURRENT APPLICATION NUMBER: US/09/014,416
CURRENT FILING DATE: 1998-01-27
EARLIER APPLICATION NUMBER: US 60/053,062
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 65
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 6
LENGTH: 9599
TYPE: DNA
ORGANISM: Hepatitis C virus
US-09-014-416-6
```

```
Query Match      87.6%; Score 730.6; DB 3; Length 9599;
Best Local Similarity 93.9%; Pred. No. 1.1e-212;
Matches 782; Conservative 0; Mismatches 49; Indels 2; Gaps 2;
```

```
QY 1 AAAAAAAAAAAGTAACCAACCTCCGCGCAGAGAGTGTCCCGGGCGGTGT 60
DB 366 AAAAAAAAAAAGTAACCAACCTCCGCGCAGAGAGTGTCCCGGGCGGTGT 425
QY 61 CAGATCGTGTGTGAGTTTAACTGTGTGCGCGCAGAGGCGCCAGTTGGGTGTGCGCGG 120
DB 426 CAGATCGTGTGTGAGTTTAACTGTGTGCGCGCAGAGGCGCCAGTTGGGTGTGCGCGG 485
QY 121 ACTAGGAAGACTTCCGAGCGGTGTGCAACTGTGTGAAGAGCGACAACCTATCCCAAGGCT 180
DB 486 ACTAGGAAGACTTCCGAGCGGTGTGCAACTGTGTGAAGAGCGACAACCTATCCCAAGGCT 545
QY 181 CCGCAGCGCGAGGCGAGGCGCTGCGGCTCAGCCCGGGTACCTTGGCCCTTATGCGCAGC 240
DB 546 CCGCAGCGCGAGGCGAGGCGCTGCGGCTCAGCCCGGGTACCTTGGCCCTTATGCGCAGC 605
QY 241 GAGGCGATGGGCGAGGCGATGGCTCTGTCAACCCCGTGGCTCCCGGCTATGGGGGCT 300
DB 606 GAGGCGATGGGCGAGGCGATGGCTCTGTCAACCCCGTGGCTCCCGGCTATGGGGGCT 665
QY 301 CCGACTGACCCCGGCGGTAGTGGGTAAATTGGGTAAAGTCAATGATCCTTACATGTC 360
DB 666 CCGACTGACCCCGGCGGTAGTGGGTAAATTGGGTAAAGTCAATGATCCTTACATGTC 725
QY 361 GCGTTCGCGCATCTTCATAGGGGTACATTCGCTGCTGCGGCTCCCTTA-GGGGGCGCTGC 419
DB 726 GCGTTCGCGCAT-TCATATGGGTACATTCGCTGCTGCGGCTCCCTTAAGGGGGCGCTGC 784
QY 420 CAGGCGCTTGCGCATGGCTCGGCGTTCGAGAGACGCGGTGAATATATGCAACAGGAA 479
DB 785 CAGGCGCTTGCGCATGGCTCGGCGTTCGAGAGACGCGGTGAATATATGCAACAGGAA 844
QY 480 TTATACCGGTTGCTCTTTCTATCTTCTCTTGGCTTTGCTGCTGTTTGAACATTC 539
DB 845 CTTCGCGGTTGCTCTTTCTATCTTCTCTTGGCTTTGCTGCTGTTTGAACATTC 904
```

QY 540 AGCTTCGGTATGAGTGCAGACGCTGTCGGGATCTACATGTCACGAAGATTGCTC 599  
| | | | |  
DB 905 AGCTTCGGTATGAGTGCAGACGCTGTCGGGATCTACATGTCACGAAGATTGCTC 964  
| | | | |  
QY 600 CAACCTCAAGCATGCTGTACGAGACAGCGGACATGATCATGACACACCCCGGGTGTGTGC 659  
| | | | |  
DB 965 CAACCTCAAGCATGCTGTACGAGACAGCGGACATGATCATGATCATCTCCGGGTGTGTGC 1024  
| | | | |  
QY 660 CTGTGTCCGGAGGGGTAATTCCTCCGCTGCTGGTAGGCGTCACTCCACGCTCGCGGC 719  
| | | | |  
DB 1025 CTGTGTACGAGAGGGTAAGAGCTCCCGTGTCTGGGTAGGCTCACTCCACGCTCGCGGC 1084  
| | | | |  
QY 720 CAAGACGCGACGATCCCACTGCGCACAATACGACGCGCATGATGATGCTGTTGGGAC 779  
| | | | |  
DB 1085 CAGGATATGCGACGCTCCCACTACGACATACGACGCGCATGATGATGCTGTTGGGAC 1144  
| | | | |  
QY 780 GGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 832  
| | | | |  
DB 1145 GGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197  
| | | | |

RESULT 10  
US-09-539-601-1  
; Sequence 1, Application US/09539601C  
; Patent No. 6630343

; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11076  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct I389/Core-3'/wt  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1342)..(1193)  
; OTHER INFORMATION: hepatitis C virus core - neomycin  
; OTHER INFORMATION: phosphotransferase fusion protein  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (1202)..(1812)  
; OTHER INFORMATION: internal ribosome entry site from  
; OTHER INFORMATION: encephalomyocarditis virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1813)..(10845)  
; OTHER INFORMATION: hepatitis C virus polyprotein from core to  
; OTHER INFORMATION: nonstructural protein NS5B; parental sequence  
; OTHER INFORMATION: without cell culture-adaptive mutations  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: (10846)..(11076)  
US-09-539-601-1

Query Match 87.6%; Score 730.6; DB 4; Length 11076;  
Best Local Similarity 93.9%; Pred. No. 1.2e-212;  
Matches 782; Conservative 0; Mismatches 49; Indels 2; Gaps 2;

QY 1 AGAAAAACCAAGTAACGACCAACCTCCGCGACAGGAGTACGTTCCCGGGCGGTGT 60  
| | | | |  
DB 1837 AAAAAACCAAGTAACGACCAACCTCCGCGACAGGAGTACGTTCCCGGGCGGTGT 1896  
| | | | |  
QY 61 CAGATGTTGTGTGAGTTTACTGTTGCGCGAGGGGCCCAAGTTGGGTGTGCGCGG 120  
| | | | |

DB 1897 CAGATGTTGTGTGAGTTTACTGTTGCGCGAGGGGCCCAAGTTGGGTGTGCGCGG 1956  
| | | | |  
QY 121 ACTAGGAAGACTCTCCGAGGAGGTCGCAACCTGTGGAGAGGACAACTATCCCAAGGCT 180  
| | | | |  
DB 1957 ACTAGGAAGACTCTCCGAGGAGGTCGCAACCTGTGGAGAGGACAACTATCCCAAGGCT 2016  
| | | | |  
QY 181 CGCCAGCCCGAGAGGAGGCTGTGGCTCAGCCCGGATACCTTGGCCCTCTATGGCAAC 240  
| | | | |  
DB 2017 CGCCAGCCCGAGAGGAGGCTGTGGCTCAGCCCGGATACCTTGGCCCTCTATGGCAAT 2076  
| | | | |  
QY 241 GAGGCGATGGGCTGGGCGATGATGCTCTGTCAACCCGTGGCTCCGGCTATGTTGGGCG 300  
| | | | |  
DB 2077 GAGGCGATGGGCTGGGCGATGATGCTCTGTCAACCCGTGGCTCCGGCTATGTTGGGCG 2136  
| | | | |  
QY 301 CCCACTGACCCCGCGAGGATCGGTAATTTGGGTAAAGTATGATATACCTTCATGCG 360  
| | | | |  
DB 2137 CCAAGGACCCCGCGAGGATCGGTAATTTGGGTAAAGTATGATATACCTTCATGCG 2196  
| | | | |  
QY 361 GGCTTCGCGGACTCTCATGAGGATACATTCGCTCGCGGCTCCCTTA-GGGGCGCTGC 419  
| | | | |  
DB 2197 GGCTTCGCGGACTCTCATGAGGATACATTCGCTCGCGGCTCCCTTAAGGGGCGCTGC 2255  
| | | | |  
QY 420 CAGGCGCTTCGCGGACTGCGCTCGCGGCTTCGAGAGACGCGCTGAATATGCAACAGGAA 479  
| | | | |  
DB 2256 CAGGCGCTTCGCGGACTGCGCTCGCGGCTTCGAGAGACGCGCTGAATATGCAACAGGAA 2315  
| | | | |  
QY 480 TTTACCGGCTGCTCTCTCATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539  
| | | | |  
DB 2316 TTTACCGGCTGCTCTCTCATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2375  
| | | | |  
QY 540 AGCTTCGGTATGAGTGCAGACGCTGTCGGGATCTACATGTCACGAAGATTGCTC 599  
| | | | |  
DB 2376 AGCTTCGGTATGAGTGCAGACGCTGTCGGGATCTACATGTCACGAAGATTGCTC 2435  
| | | | |  
QY 600 CAACCTCAAGCATGCTGTACGAGACAGCGGACATGATCATGACACACCCCGGGTGTGTGC 659  
| | | | |  
DB 2436 CAACCTCAAGCATGCTGTACGAGACAGCGGACATGATCATGATCATCTCCCGGGTGTGTGC 2495  
| | | | |  
QY 660 CTGTGTCCGGAGGGGTAATTCCTCCGCTGCTGGTAGGCGTCACTCCACGCTCGCGGC 719  
| | | | |  
DB 2496 CTGTGTCCGGAGGGGTAATTCCTCCGCTGCTGGTAGGCGTCACTCCACGCTCGCGGC 2555  
| | | | |  
QY 720 CAAGACGCGACGATCCCACTGCGCACAATACGACGCGCATGATGATGCTGTTGGGAC 779  
| | | | |  
DB 2556 CAAGACGCGACGATCCCACTGCGCACAATACGACGCGCATGATGATGCTGTTGGGAC 2615  
| | | | |  
QY 780 GGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 832  
| | | | |  
DB 2616 GGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2668  
| | | | |

RESULT 11  
US-09-539-601-19  
; Sequence 19, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 11076  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)

```

; OTHER INFORMATION: construct I389/Core-3'/9-13F
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1302)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polypeptide from core to
; OTHER INFORMATION: nonstructural protein NS5B, carries cell
; OTHER INFORMATION: culture-adaptive mutations from clone 9-13F
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: (10846)..(11076)
US-09-539-601-19

Query Match      87.6%; Score 730.6; DB 4; Length 11076;
Best Local Similarity 93.9%; Pred. No. 1.2e-212;
Matches 782; Conservative 0; Mismatches 49; Indels 2; Gaps 2;

QY 1 AGAAAAACCAAGTAACCAACCTCCGCCACAGAGCTCAGTTCCCGGGCGGTGT 60
DB 1837 AGAAAAACCAAGTAACCAACCTCCGCCACAGAGCTCAGTTCCCGGGCGGTGT 1896
QY 61 CAGATCGTGTGGAGTTTACTGTTGCGCGAGAGGGGCCCGAGTTGGGTGCGCGG 120
DB 1897 CAGATCGTGTGGAGTTTACTGTTGCGCGAGAGGGGCCCGAGTTGGGTGCGCGG 1956
QY 121 ACTAGAAAGACTTCCGAGGGGTGCAACCTGTGGAAGGAGCAACCTATCCCAAGCT 180
DB 1957 ACTAGAAAGACTTCCGAGGGGTGCAACCTGTGGAAGGAGCAACCTATCCCAAGCT 2016
QY 181 CGCAGACCCGAGGGGCGAGGGGCTGAGCCCGGGTACCTTGGCCCTTATGGCAAC 240
DB 2017 CGCAGACCCGAGGGGCGAGGGGCTGAGCCCGGGTACCTTGGCCCTTATGGCAAC 2076
QY 241 GAGGGGATGGGGGTGAGGAGGAGGCTCTGTCACCCCGTGGCTCCGAGCTAGTGGGG 300
DB 2077 GAGGGGATGGGGGTGAGGAGGAGGCTCTGTCACCCCGTGGCTCCGAGCTAGTGGGG 2136
QY 301 CCCACTGACCCCGGGGTGAGGAGGCTGATTTGGGTAAGTCAATGATCCCTACATGC 360
DB 2137 CCCACTGACCCCGGGGTGAGGAGGCTGATTTGGGTAAGTCAATGATCCCTACATGC 2196
QY 361 GGGTTGGCGGACTCTCATGGGGTAACATTCGCTGTCGGGCTCCCTTA-GGGGGCGTGC 419
DB 2197 GGGTTGGCGGACTCTCATGGGGTAACATTCGCTGTCGGGCTCCCTTA-GGGGGCGTGC 2255
QY 420 CAGGGGCTTGGGCGATGGGCTCGGGGTTCTGAGAGAGCGGGTAACATGCAACAGAGAA 479
DB 2256 CAGGGGCTTGGGCGATGGGCTCGGGGTTCTGAGAGAGCGGGTAACATGCAACAGAGAA 2315
QY 480 TTTACCGGTTGCTCTTCTCTATCTTCTCTTGGCTTGGCTGCTGTTTGAACATTC 539
DB 2316 TTTACCGGTTGCTCTTCTCTATCTTCTCTTGGCTTGGCTGCTGTTTGAACATTC 2375
QY 540 AGTTTCGGCTTATGAAGTGGCAAGTGTCCGGAGTCTACATGTCACGAACATTTGCTC 599
DB 2376 AGTTTCGGCTTATGAAGTGGCAAGTGTCCGGAGTCTACATGTCACGAACATTTGCTC 2435
QY 600 CAACTCAAGCATGTGTAGAGAGAGAGGAGCATGATCATGACACCCCGGGGTGTGCC 659
DB 2436 CAACTCAAGCATGTGTAGAGAGAGAGGAGCATGATCATGACACCCCGGGGTGTGCC 2495
QY 660 CTGTGTCCGGAGGAGTAATTCCTCCGCTGCTGGGTAGCGCTCACTCCAGCTCGCGC 719
DB 2496 CTGTGTCCGGAGGAGTAATTCCTCCGCTGCTGGGTAGCGCTCACTCCAGCTCGCGC 2555
```

```

QY 720 CAAGAGCGCAGATCCCACTGCGACAAATACAGAGCCAGTCGATTTGCTGTTGGGC 779
DB 2556 CAGAAACGCTAAGCTGCTCCCACTACAGACATACAGAGCCAGTCGATTTGCTGTTGGGC 2615
QY 780 GGGTGCCTTCTGCTGCGCTATGATGCTGGGGAGATCTTCCGAGATCTGTTTCC 832
DB 2616 GGGTGCCTTCTGCTGCGCTATGATGCTGGGGAGATCTTCCGAGATCTGTTTCC 2668

RESULT 12
US-09-539-601-25
; Sequence 25, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartschlag, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199-04-03
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 25
LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/5.1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polypeptide from core to
; OTHER INFORMATION: nonstructural protein NS5B, carries cell
; OTHER INFORMATION: culture-adaptive mutations of clone 5.1
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: (10846)..(11076)
US-09-539-601-25

Query Match      87.6%; Score 730.6; DB 4; Length 11076;
Best Local Similarity 93.9%; Pred. No. 1.2e-212;
Matches 782; Conservative 0; Mismatches 49; Indels 2; Gaps 2;

QY 1 AGAAAAACCAAGTAACCAACCTCCGCCACAGAGCTCAGTTCCCGGGCGGTGT 60
DB 1837 AGAAAAACCAAGTAACCAACCTCCGCCACAGAGCTCAGTTCCCGGGCGGTGT 1896
QY 61 CAGATCGTGTGGAGTTTACTGTTGCGCGAGAGGGGCCCGAGTTGGGTGCGCGG 120
DB 1897 CAGATCGTGTGGAGTTTACTGTTGCGCGAGAGGGGCCCGAGTTGGGTGCGCGG 1956
QY 121 ACTAGAAAGACTTCCGAGGGGTGCAACCTGTGGAAGGAGCAACCTATCCCAAGCT 180
DB 1957 ACTAGAAAGACTTCCGAGGGGTGCAACCTGTGGAAGGAGCAACCTATCCCAAGCT 2016
QY 181 CGCAGACCCGAGGGGCGAGGGGCTGAGCCCGGGTACCTTGGCCCTTATGGCAAC 240
DB 2017 CGCAGACCCGAGGGGCGAGGGGCTGAGCCCGGGTACCTTGGCCCTTATGGCAAC 2076
```



US-08-462-195-1  
; Sequence 1, Application US/08462195  
; Patent No. 5789544  
; GENERAL INFORMATION:  
; APPLICANT: MIYAMURA, TATSUO  
; APPLICANT: SAITO, IZUMU  
; APPLICANT: MATSURA, YOSHIHARU  
; APPLICANT: HONDA, YOSHIKAZU  
; APPLICANT: SEKI, MAKOTO  
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462,195  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 514  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,303  
; FILING DATE: 22-MAY-1995  
; APPLICATION NUMBER: US 08/074,584  
; FILING DATE: 11-JUN-1993  
; APPLICATION NUMBER: JP 152487/1992  
; FILING DATE: 11-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5789544man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 4169-003-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ. ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1037 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Hepatitis C virus  
; IMMEDIATE SOURCE:  
; CLONE: pUC010  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 17..1036  
US-08-462-195-1

Query Match 86.3%; Score 719.4; DB 1; Length 1037;  
Best Local Similarity 93.0%; Pred. No. 9.7e-210;  
Matches 775; Conservative 0; Mismatches 56; Indels 2; Gaps 2;

QY 1 AGAAAAACCAACGTAACCAACCTCCGCAAGGAGCTTCCGGCGGTGT 60  
DB 41 AGAAAAACCAACGTAACCAACCTCCGCAAGGAGCTTCCGGCGGTGT 100  
QY 61 CAGATGCTTGTGAGTTTACCTGTTGCGCGGAGGGGCCCAAGTTGGTGGCGG 120  
DB 101 CAGATGCTTGTGAGTTTACCTGTTGCGCGGAGGGGCCCAAGTTGGTGGCGG 160

QY 121 ACTAGAGACACTCCGAGGCGTGCACACTCTGTGAAAGGAGCAACCTATCCCAAGGCT 180  
DB 161 ACTAGAGACACTCCGAGGCGTGCACACTCTGTGAAAGGAGCAACCTATCCCAAGGCT 220  
QY 181 CGCCAGCCCGAGGCGAGGCTGAGGCTCAGCCCGGATACCTTTGGCCCTTATGGCAAC 240  
DB 221 CGCCAGCCCGAGGCGAGGCTGAGGCTCAGCCCGGATACCTTTGGCCCTTATGGCAAC 280  
QY 241 GAGGGCATGGGGTGGGAGGATGGCTCCGTGACCCCGTGGCTCCGGGCTTATGGGGC 300  
DB 281 GAGGGCATGGGGTGGGAGGATGGCTCCGTGACCCCGTGGCTCCGGGCTTATGGGGC 340  
QY 301 CCCTAGACCCCGGAGGAGTGGCTGAGGCTTAAATTTGGTAAATGATCACTTACATGC 360  
DB 341 CCTATGACCCCGGAGGAGTGGCTGAGGCTTAAATTTGGTAAATGATCACTTACATGC 400  
QY 361 GGCTTGGCGGACTCTGATGGGATGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419  
DB 401 GGCTTGGCGGACTCTGATGGGATGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459  
QY 420 CAGGGCCCTGAGGAGGAGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479  
DB 460 CAGGGCCCTGAGGAGGAGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 519  
QY 480 TTACCCGAGTGGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 539  
DB 520 TTACCCGAGTGGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 579  
QY 540 AGCTTCGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599  
DB 580 AGCTTCGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 639  
QY 600 CAAGTCAAGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659  
DB 640 CAAGTCAAGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 699  
QY 660 CTGCTGCGGAGGAGGATTCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 719  
DB 700 CTGCTGCGGAGGAGGATTCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759  
QY 720 CAGAGGCGCAGGATCCCACTGCGACAAATACGACGCAAGTGTGCTGCTGCTGCTGCTG 779  
DB 760 CAGAGGCGCAGGATCCCACTGCGACAAATACGACGCAAGTGTGCTGCTGCTGCTGCTG 819  
QY 780 GGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 832  
DB 820 GGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 872

RESULT 15  
US-08-636-883-1  
; Sequence 1, Application US/08636883  
; Patent No. 5830691  
; GENERAL INFORMATION:  
; APPLICANT: MIYAMURA, TATSUO  
; APPLICANT: SAITO, IZUMU  
; APPLICANT: MATSURA, YOSHIHARU  
; APPLICANT: HONDA, YOSHIKAZU  
; APPLICANT: SEKI, MAKOTO  
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/636,883  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,303  
FILING DATE:  
APPLICATION NUMBER: US 08/074,584  
FILING DATE: 11-JUN-1993  
APPLICATION NUMBER: JP 152487/1992  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: OBIOM, NO. 5830691man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 4169-003-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1037 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Hepatitis C virus  
IMMEDIATE SOURCE:  
CLONE: pUC010  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 17..1036  
US-08-636-883-1

Query Match 86.3%; Score 719.4; DB 2; Length 1037;

Best Local Similarity 93.0%; Pred. No. 9,7e-210; Mismatches 775; Conservative 0; Mismatches 56; Indels 2; Gaps 2;

QY 1 AGAAAAACCAACGTAACCAACCTCCGCCACAGAGAGTTCCTCCGGCGGTGGT 60  
DB 41 AGAAAAACCAACGTAACCAACCTCCGCCACAGAGAGTTCCTCCGGCGGTGGT 100  
QY 61 CAGATCGTGTGGAGTTAAGTTCCTGTCGCGCGAGGGGCCCCAGGTGGGTGTGCGCG 120  
DB 101 CAGATCGTGTGGAGTTAAGTTCCTGTCGCGCGAGGGGCCCCAGGTGGGTGTGCGCG 160  
QY 121 ACTAGGAAGACTTCGAGGGGTGCAACTCTGTGGAAGGCGACAACCTATCCCAAGGCT 180  
DB 161 ACTAGGAAGACTTCGAGGGGTGCAACTCTGTGGAAGGCGACAACCTATCCCAAGGCT 220  
QY 181 GCGCAAGCCGAGGCGAGGGCTGAGCCCGGGTACCTTGGCCCTCTATGGCAAC 240  
DB 221 GCGCGGCGCGAGGCGAGGAGCTGAGGCTCAGCCGAGGTATCCTTGGCCCTCTATGGCAAT 280  
QY 241 GAGGGCATGGGGTGGGCAAGAGTGTCTCTGTCACTCCGTCGCTCCGGGCTTGTGGGGC 300  
DB 281 GAGGGCTTGGGGTGGGCAAGAGTGTCTCTGTCACTCCGTCGCTCCGGGCTTGTGGGGC 340  
QY 301 CCACGTGACCCCGGGGTGAGTGGGTATTTGGGTAAGTCATGATGATCCCTCATGTC 360  
DB 341 CCTAATGACCCCGGGGTGAGTGGGTATTTGGGTAAGTCATGATGATCCCTCATGATGC 400  
QY 361 GGCTTCGCGACTTCATGAGGGGTACATTCGCTCGTGGCGGCTCCCTTA-GGGCGGCTGC 419  
DB 401 GGCTTCGCGGAC-CTCATGGGGTACATTCGCTCGTGGCGGCTCCCTTAAGGGGGGCTGC 459  
QY 420 CAGGGCCCTGGCGCATGCGGTCTCGGGTTCTGAGAGACGGCGTGAATATGCAACAGGANA 479  
DB 460 CAGGGCCCTGGCGCATGCGGTCTCGGGTTCTGAGAGACGGCGTGAATATGCAACAGGANA 519

QY 480 TTTAACCGGTTGCTCTTCTCTATCTCTCTTGGCTTGTGCTGTCTGTGACCATTC 539  
DB 520 TTTCGCGGTTGCTCTTCTCTATCTCTCTTGTAGCTGTGCTGTCTGTGACCATTC 579  
QY 540 AGCTTCGCTTATGAGTGGCGCAACGTGTCCGGGATCTAACATGTCAAGAAAGATTGCTC 599  
DB 580 AGCTTCGCTTATGAGTGGCGCAACGTGTCCGGGATCTAACATGTCAAGAAAGATTGCTC 639  
QY 600 CAACGTGAGCATCGTGTGAGAGACAGCGGAGATGATGACACACCCCGGGGTGTGTC 659  
DB 640 CAACGTGAGCATGTGTGTGAGAGCGGAGATGATGACATGACATGCCCCGGGTGTGTC 699  
QY 660 CTGTGTCGAGAGAGGTAAATTCCTCCGCTGTGAGTACGCTCACTCCACGCTGCGAGC 719  
DB 700 CTGCGTTGGGAGAAATTCCTCCGCTGTGAGTACGCTCACTCCACGCTGCGAGC 759  
QY 720 CAGGACCGCAGCATCCCACTGCGCAATACAGACGCAAGTGTGATTTGCTGTGGGAC 779  
DB 760 CAGGATGCGAGCGTCCCACTACGACATTAAGACGCGACGTCGACTGTGCTGTGGGAC 819  
QY 780 GGCTGCTTCTGTCGCTATGATGACGTGGGGATCTGCGGATGTGTTTC 832  
DB 820 GGCTGCTTCTGCTGCTATGATGACGTGGGGATCTGCGGATGTGTTTC 872

Search completed: February 21, 2005, 04:09:01  
Job time : 198 secs





QY 61 ACTGCACAAATAGACAGCCACGTCGATTGCTGGGGGGGCTTCTGCTCGCT 120  
|  
|  
|  
Db 61 ACTGCACAAATAGACAGCCACGTCGATTGCTGGGGGGGCTTCTGCTCGCT 120  
|  
|  
|  
QY 121 ATGTACGTGGGGGATCTCTGCGGATCTGTTTCTCTGCTCTGATGTTACCTTCTG 180  
|  
|  
|  
|  
Db 121 ATGTACGTGGGGGATCTCTGCGGATCTGTTTCTCTGCTCTGATGTTACCTTCTG 180  
|  
|  
|  
QY 181 CCTGCGGACATAGACAGGATGACAGGACTGCAATTGTTCAATCTATCCCGGCGCATGCA 240  
|  
|  
|  
|  
Db 181 CCTGCGGACATAGACAGGATGACAGGACTGCAATTGTTCAATCTATCCCGGCGCATGCA 240  
|  
|  
|  
QY 241 GGTACCGGATGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
|  
|  
|  
|  
Db 241 GGTACCGGATGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
|  
|  
|  
QY 301 TCGCAGCTACTCCGATCCCAAGCTGTGTGCACTGGTGGGGGGGCGCCACTGGGGA 360  
|  
|  
|  
|  
Db 301 TCGCAGCTACTCCGATCCCAAGCTGTGTGCACTGGTGGGGGGGCGCCACTGGGGA 360  
|  
|  
|  
QY 361 GTCCTGGCGGGGCTTGGCTACTATTCATGATGATGATGATGATGATGATGATGATGATGAT 420  
|  
|  
|  
|  
Db 361 GTCCTGGCGGGGCTTGGCTACTATTCATGATGATGATGATGATGATGATGATGATGATGAT 420  
|  
|  
|  
QY 421 ATGCTACTCTTTGCGCGGCTTGAAGGGAACCTTACACGACAGGGGGGACACACGCGCG 480  
|  
|  
|  
|  
Db 421 ATGCTACTCTTTGCGCGGCTTGAAGGGAACCTTACACGACAGGGGGGACACACGCGCG 480  
|  
|  
|  
QY 481 GCCGCCACAGGGGCTTACATCTCTCTTCAACCTGGGGGCGGCTCAAGAAATCCAGCTTGT 540  
|  
|  
|  
|  
Db 481 GCCGCCACAGGGGCTTACATCTCTCTTCAACCTGGGGGCGGCTCAAGAAATCCAGCTTGT 540  
|  
|  
|  
QY 541 AACACCAAGGGGCTGAGCATCAACAGAACTGCTTGAATGCAATGATGATGATGATGATGAT 600  
|  
|  
|  
|  
Db 541 AACACCAAGGGGCTGAGCATCAACAGAACTGCTTGAATGCAATGATGATGATGATGATGAT 600  
|  
|  
|  
QY 601 ACTGGGTTCTTTCGCGGCTGTTCTACACGACAGGTTCAATCGTCCGATCTCAGAG 660  
|  
|  
|  
|  
Db 601 ACTGGGTTCTTTCGCGGCTGTTCTACACGACAGGTTCAATCGTCCGATCTCAGAG 660  
|  
|  
|  
QY 661 CGCATGGGCGAGCTGCGGCCCATTTGACAGATGATGATGATGATGATGATGATGATGAT 720  
|  
|  
|  
|  
Db 661 CGCATGGGCGAGCTGCGGCCCATTTGACAGATGATGATGATGATGATGATGATGATGAT 720  
|  
|  
|  
QY 721 AATGAGTCCACGAGCTTGAACAGAGGCTTATTTGCTGAGCACTACCAACCTGAT 780  
|  
|  
|  
|  
Db 721 AATGAGTCCACGAGCTTGAACAGAGGCTTATTTGCTGAGCACTACCAACCTGAT 780  
|  
|  
|  
QY 781 GGTATGCTGCGCGGCTTGAACAGAGGCTTATTTGCTGAGCACTACCAACCTGAT 840  
|  
|  
|  
|  
Db 781 GGTATGCTGCGCGGCTTGAACAGAGGCTTATTTGCTGAGCACTACCAACCTGAT 840  
|  
|  
|  
QY 841 GTGGTGGGAGCAACCGATGTTTGGGCGCCCTTACGTAAGATGGGGTGAATGAGACG 900  
|  
|  
|  
|  
Db 841 GTGGTGGGAGCAACCGATGTTTGGGCGCCCTTACGTAAGATGGGGTGAATGAGACG 900  
|  
|  
|  
QY 901 GATGCTGCTGTTCTCAACAGAGGCGGCGCCCAAGGCTGATGATGATGATGATGATGATGAT 960  
|  
|  
|  
|  
Db 901 GATGCTGCTGTTCTCAACAGAGGCGGCGCCCAAGGCTGATGATGATGATGATGATGATGAT 960  
|  
|  
|  
QY 961 ATGAATAGCAACCGGTTCAACAGAGCGTGTGGGGGGGCGCCGCTGCAACATCGGGGGGCT 1020  
|  
|  
|  
|  
Db 961 ATGAATAGCAACCGGTTCAACAGAGCGTGTGGGGGGGCGCCGCTGCAACATCGGGGGGCT 1020  
|  
|  
|  
QY 1021 GGGCAACAACCTTTGATCTGCGCCACGAGCTGTTCTGGAAGCATCCGAGGCACTTAC 1080  
|  
|  
|  
|  
Db 1021 GGGCAACAACCTTTGATCTGCGCCACGAGCTGTTCTGGAAGCATCCGAGGCACTTAC 1080  
|  
|  
|  
QY 1081 ACCAAATGCGGTTGGGGGCTTGGTTG 1107  
|  
|  
|  
|  
Db 1081 ACCAAATGCGGTTGGGGGCTTGGTTG 1107  
|  
|  
|

RESULT 2  
A32200  
LOCUS  
DEFINITION NANBH PT polypeptide fragment.  
ACCESSION A32200  
VERSION A32200.1 GI:1926558  
KEYWORDS  
SOURCE  
ORGANISM Non-A, non-B hepatitis virus  
Non-A, non-B hepatitis virus  
Virus; unclassified viruses.  
REFERENCE  
1. Patent: FR 2655990-A 19 21-JUN-1991;  
Location/Qualifiers  
FEATURES  
source  
1..1107  
/organism="Non-A, non-B hepatitis virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:12440"  
/clone="136/155"  
<1..>1107  
CDS  
1..1107  
/codon\_start=1  
/product="partial NANBH-PT polypeptide"  
/protein\_id="CA02092.1"  
/db\_xref="GI:1926559"  
/translation="SRRCWALPTLAKDASIPYATIRRHVLLVGAAPCSAMTV  
DLGSVFLVSLPTSPRRHOTVODNCISLYGHVSGHRMAMDMNMSPTALVYSO  
LIRIQAIVDMVAGAHVGLAGLAVSMVGNNAKYLIVMLFAGVDGEPYTGALVSG  
AAHGTSLETPPBAOKIQLVNTHNGSHIRPLNCHDSIQTFPLALPYTHRNASGC  
SRMASCRPIDQDPQWGPITYNESHGIDQRYCMYAPQGIYALVDCGVYCFY  
PSVVGTTDRGAPTYRWGENETDVLNINRPPRGNNFGCTMNNSTGFTYTCGGPP  
CNIGGVNNTLICPDCEFRKHPEATYTKGSGPMW"  
ORIGIN  
Query Match 100.0%; Score 1107; DB 6; Length 1107;  
Best Local Similarity 100.0%; Pred. No. 1.1e-232; Indels 0; Gaps 0;  
Matches 1107; Conservative 0; Mismatches 0;  
QY 1 TCTCTCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCGCAAGGACGACATCCCC 60  
|  
|  
|  
|  
Db 1 TCTCTCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCGCAAGGACGACATCCCC 60  
|  
|  
|  
QY 61 ACTGCACAAATAGACAGCCACGTCGATTGCTGTTGGGGGCGGCTTCTGCTCGCT 120  
|  
|  
|  
|  
Db 61 ACTGCACAAATAGACAGCCACGTCGATTGCTGTTGGGGGCGGCTTCTGCTCGCT 120  
|  
|  
|  
QY 121 ATGTACGTGGGGGATCTCTGCGGATCTGTTTCTCTGCTCTGACGTTTCACTTCTG 180  
|  
|  
|  
|  
Db 121 ATGTACGTGGGGGATCTCTGCGGATCTGTTTCTCTGCTCTGACGTTTCACTTCTG 180  
|  
|  
|  
QY 181 CCTGCGGACATAGACAGGATGACAGGACTGCAATTGTTCAATCTATCCCGGCGCATGCA 240  
|  
|  
|  
|  
Db 181 CCTGCGGACATAGACAGGATGACAGGACTGCAATTGTTCAATCTATCCCGGCGCATGCA 240  
|  
|  
|  
QY 241 GGTACCGGATGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
|  
|  
|  
|  
Db 241 GGTACCGGATGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
|  
|  
|  
QY 301 TCGCAGCTACTCCGATCCCAAGCTGTGTGCACTGGTGGGGGGGCGCCACTGGGGA 360  
|  
|  
|  
|  
Db 301 TCGCAGCTACTCCGATCCCAAGCTGTGTGCACTGGTGGGGGGGCGCCACTGGGGA 360  
|  
|  
|  
QY 361 GTCCTGGCGGGGCTTGGCTACTATTCATGATGATGATGATGATGATGATGATGATGATGAT 420  
|  
|  
|  
|  
Db 361 GTCCTGGCGGGGCTTGGCTACTATTCATGATGATGATGATGATGATGATGATGATGATGAT 420  
|  
|  
|  
QY 421 ATGCTACTCTTTGCGCGGCTTGAAGGGAACCTTACACGACAGGGGGGACACACGCGCG 480  
|  
|  
|  
|  
Db 421 ATGCTACTCTTTGCGCGGCTTGAAGGGAACCTTACACGACAGGGGGGACACACGCGCG 480  
|  
|  
|  
QY 481 GCCGCCACAGGGGCTTACATCTCTCTTCAACCTGGGGGCGGCTCAAGAAATCCAGCTTGT 540  
|  
|  
|  
|  
Db 481 GCCGCCACAGGGGCTTACATCTCTCTTCAACCTGGGGGCGGCTCAAGAAATCCAGCTTGT 540  
|  
|  
|

QY	541	AACACCAACGGGACGTGGCACTACAAAGAACTGGCTTGAATGGAAATGACCTCCCTCAA	600
Db	541	AACACCAACGGGACGTGGCACTACAAAGAACTGGCTTGAATGGAAATGACCTCCCTCAA	600
QY	601	ACTGGGTTCTTGGCCGCGCTGTTCTACAACGACAGGTTCAAATGCGTCCGGAATGCTCAGAG	660
Db	601	ACTGGGTTCTTGGCCGCGCTGTTCTACAACGACAGGTTCAAATGCGTCCGGAATGCTCAGAG	660
QY	661	CGCATGGCCAGCTGCGCCCATTTGACCAAGTTCCATCAAGGGGTGGGGTCCCATCACTTAT	720
Db	661	CGCATGGCCAGCTGCGCCCATTTGACCAAGTTCCATCAAGGGGTGGGGTCCCATCACTTAT	720
QY	721	AATGAGTCCCAACGGCTTGAGCAACAGGGCCCTAATGTGTCGCACTACAGCACTCAACCGGT	780
Db	721	AATGAGTCCCAACGGCTTGAGCAACAGGGCCCTAATGTGTCGCACTAAGCACTCAACCGGT	780
QY	781	GGTATTCGTGCCCGCGCTTGCAGGTGTGTGGCCAGTGTACTGTCTTTCACATCCAAAGCCCTGT	840
Db	781	GGTATTCGTGCCCGCGCTTGCAGGTGTGTGGCCAGTGTACTGTCTTTCACATCCAAAGCCCTGT	840
QY	841	GTGCTGGGGAGCAGCCGATCGTTTGGGCGCCCTCACTACAGATGGGGTGAGAAATGAGACG	900
Db	841	GTGCTGGGGAGCAGCCGATCGTTTGGGCGCCCTCACTACAGATGGGGTGAGAAATGAGACG	900
QY	901	GACGTCGTCGCTTTCACAACAACGCGGGCGGCAACGGGGCAACTGTTCGGCTGTACATGG	960
Db	901	GACGTCGTCGCTTTCACAACAACGCGGGCGGCAACGGGGCAACTGTTCGGCTGTACATGG	960
QY	961	ATGAATAGACACCGGGTTCACCAAGAAGCTGTGTGGGGCCCCCGGTGCAACATCGGGGGGGTTC	1020
Db	961	ATGAATAGACACCGGGTTCACCAAGAAGCTGTGTGGGGCCCCCGGTGCAACATCGGGGGGGTTC	1020
QY	1021	GGCAACAACACTTTGATCTGCCCCACGGAAGCTCTTCGGAAGCAATCCCGAAGGCCACTTAC	1080
Db	1021	GGCAACAACACTTTGATCTGCCCCACGGAAGCTCTTCGGAAGCAATCCCGAAGGCCACTTAC	1080
QY	1081	ACCAAAATCCGGTTCGGGGCTTTGGTTG 1107	
Db	1081	ACCAAAATCCGGTTCGGGGCTTTGGTTG 1107	

RESULT	3				
LOCUS	AR144048				
DEFINITION	Sequence 19 from patent US 6210675.	1107 bp	DNA	linear	PAT 08-AUG-2001
ACCESSION	AR144048				
VERSION	AR144048.1	GI:15105915			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1107) Highfield,P.Edmund., Rodgers,B.Colin., Tedder,R.Seton. and Barbara,J.Anthony.James. PT-NAMB hepatitis polyepitides Patent: US 6210675-A 18 03-APR-2001; Location/Qualifiers				
TITLE	JOURNAL				
FEATURES	Source				
	1..1107				
	/organism="unknown"				
	/mol_type="unassigned DNA"				
ORIGIN					

[illegible]

Db	61	ACTGCAACAATACAGACGACGACGTGCAATTTGCTCGTTGGGGGGGGTGCCTTCTGCTCGCCT	120
Qy	121	ATGTACGTGGGGGGATCTCTGGGGATCTGTTTTCTCTGCTCTCAAGCTGTTCAACTTCTCG	180
Db	121	ATGTACGTGGGGGGATCTCTGGGGATCTGTTTTCTCTGCTCTCAAGCTGTTCAACTTCTCG	180
Qy	181	CCTGCGGCAATCAGAGGGTACAGAGACTGCAAATGTTCAATCTATCCCGGCAAGTATCA	240
Db	181	CCTGCGGCAATCAGAGGGTACAGAGACTGCAAATGTTCAATCTATCCCGGCAAGTATCA	240
Qy	241	GGTACCCGCATGCGTTGGGATATGATGATGAATGAACTGTCACTTAACAGCACCCCTAGTGGTA	300
Db	241	GGTACCCGCATGCGTTGGGATATGATGATGAATGAACTGTCACTTAACAGCACCCCTAGTGGTA	300
Qy	301	TTCGAGCTACTCCGGATATCCCAAGCTGTGTTGGAATATGATGGGGGGGGCCCACTGGGGA	360
Db	301	TTCGAGCTACTCCGGATATCCCAAGCTGTGTTGGAATATGATGGGGGGGGCCCACTGGGGA	360
Qy	361	GTCCTGGCGGGCTTGGCTACTATTCATAGGTGGGGAACTGGGCTTAAGGTCTTGTTGTG	420
Db	361	GTCCTGGCGGGCTTGGCTACTATTCATAGGTGGGGAACTGGGCTTAAGGTCTTGTTGTG	420
Qy	421	ATGCTACTCTTTGCGGGCTTGACGGGGAACTTTACACGACAGGGGGGACACAGGCCGC	480
Db	421	ATGCTACTCTTTGCGGGCTTGACGGGGAACTTTACACGACAGGGGGGACACAGGCCGC	480
Qy	481	GGCGCCACGGGGCTTACATCCCTTTCAACCTGAGGGCGGGCTCAGAAAAATCAGCTTGTA	540
Db	481	GGCGCCACGGGGCTTACATCCCTTTCAACCTGAGGGCGGGCTCAGAAAAATCAGCTTGTA	540
Qy	541	AACACCAACGGCAGCTGGCAGATCAACAGAACTGCTTGAACCTGCAATGACTCCCTCCAA	600
Db	541	AACACCAACGGCAGCTGGCAGATCAACAGAACTGCTTGAACCTGCAATGACTCCCTCCAA	600
Qy	601	ACTGGGTTCTTGGCGGGCTGTTTACACGACAGGTTCAATGCGTCCGATGCTCAGAG	660
Db	601	ACTGGGTTCTTGGCGGGCTGTTTACACGACAGGTTCAATGCGATCGGATGCTCAGAG	660
Qy	661	CGCATGCGCAGCTCCGCCCATATGACAGTTCGATCAGGGGTGGGGTCCATCACTTAT	720
Db	661	CGCATGCGCAGCTCCGCCCATATGACAGTTCGATCAGGGGTGGGGTCCATCACTTAT	720
Qy	721	AATGATGCCACGGGCTTGACACAGAGGCCCTATATGCTGACCTACACCTGCTG	780
Db	721	AATGATGCCACGGGCTTGACACAGAGGCCCTATATGCTGACCTACACCTGCTG	780
Qy	781	GGTATCGTCCCGGGTTGACAGGTGTGTGGCCCAAGTATAGTTTCACTCCAAAGCCGTG	840
Db	781	GGTATCGTCCCGGGTTGACAGGTGTGTGGCCCAAGTATAGTTTCACTCCAAAGCCGTG	840
Qy	841	GTGGTGGGGAGCAGCCGATCGTTTGGGGCCCTTACGTACAGATGGGGTGAAGATGAGCG	900
Db	841	GTGGTGGGGAGCAGCCGATCGTTTGGGGCCCTTACGTACAGATGGGGTGAAGATGAGCG	900
Qy	901	GACGTGCTGCTTCTCAACAAACAGCGGGCCGACAGGGGCAACTGTGTCGTATCATGG	960
Db	901	GACGTGCTGCTTCTCAACAAACAGCGGGCCGACAGGGGCAACTGTGTCGTATCATGG	960
Qy	961	ATGAATAGCACCGGGTTCAACAAAGCGTGTGGGGGGCCCGCGTGCACATCGGGGGGGT	1020
Db	961	ATGAATAGCACCGGGTTCAACAAAGCGTGTGGGGGGCCCGCGTGCACATCGGGGGGGT	1020
Qy	1021	GGCAACAACCTTTGATCTGCCCCACGGACTGCTTCGGAAGCATCCCGAGGCCACTTAC	1080
Db	1021	GGCAACAACCTTTGATCTGCCCCACGGACTGCTTCGGAAGCATCCCGAGGCCACTTAC	1080
Qy	1081	ACCAATGCGGTTCCGGGGCTTGGTTG 1107	
Db	1081	ACCAATGCGGTTCCGGGGCTTGGTTG 1107	

RESULT 4  
A28155

A28155

LOCUS A28155 2116 bp DNA linear PAT 07-JUN-1995  
DEFINITION PT-NANBH mRNA fragment from patent GB2239245.  
ACCESSION A28155  
VERSION A28155.1 GI:1248638  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 2116)  
TITILE Post-transfusal non-A non-B hepatitis viral polypeptides  
JOURNAL Patent: GB 2239245-A 21 26-JUN-1991;  
FEATURES  
source Location/Qualifiers  
1..2116  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="contig formed by cDNA clones"  
308..>2116  
/codon\_start=1  
/transl\_table=1  
/product="PT-NANBH polyprotein N-terminus"  
/protein\_id="CAA01931.1"  
/db\_xref="GI:1248639"  
/translation="MSTNPKPORKTKRNTNRRPQDVKFPGGQIVGGVYLPRRGRPL  
GVRATKTSRSGRGRQPIPKARQPEGAWAOPQYPMPLVGNBGMWAGMLSPRG  
SRWSGPTPRRRSRNLGKVIDTLTGCPDLMGYIPLVGAPLGAHALAHGVRLD  
GVNVAATGNLPGCSFSIFILALISCLTIPASAYBVRNGIYHTNDCSNLSIYETAD  
MIMHTPGCVPCVRBNSRCVWALPTTLAKDASIPATIRRHVDLLVGAAPCSAMY  
VBDLGSVFLVSQLETFSPRRHQTVDNCSTYPGHVSGRHAWMDMMMSPTALV  
SOLRLIPOAVVMWAGAHWGLAGLAYSVWGMALVVMLLFAVIDEPYTTGGTH  
GRAHLTLSTFTGPAOKIQLVTNNGSMHINRTALNCNLSLQTFGALFLYTRFVAS  
GSEWASCRPIDQFQDQNGPIYVNSHGDORPYCMHVAPOGCIIVPLAIVCGPYVC  
FIPSPIVGTTPDRFGAPTRYRMGENTDVLINNTPRPGMWFCCTMNSGTGFKTCGS  
PENCIGVGNNTLICPTDCKRKPBAITTKCGSGPML"

ORIGIN  
Query Match 100.0%; Score 1107; DB 6; Length 2116;  
Best Local Similarity 100.0%; Pred. No. 1e-232;  
Matches 1107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCGGGCCAAAGACGCGACATCCCC 60  
DB 1010 TCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCGGGCCAAAGACGCGACATCCCC 1069  
QY 61 ACTGCCAATATAGAGCCCACTGCATTTGCTGTTGGGGCGCTGCTTCTGCTCGGCT 120  
DB 1070 ACTGCCAATATAGAGCCCACTGCATTTGCTGTTGGGGCGCTGCTTCTGCTCGGCT 1129  
QY 121 ATGACGTGGGGATCTCTCGGATCTGTTTCTGCTCTCAGCTGTTCACTTCTCG 180  
DB 1130 ATGACGTGGGGATCTCTCGGATCTGTTTCTGCTCTCAGCTGTTCACTTCTCG 1189  
QY 181 CTTGCCGACATCAGACGCTGACAGACTGCAATTTTCAATCTATCCGGCCACATCA 240  
DB 1190 CTTGCCGACATCAGACGCTGACAGACTGCAATTTTCAATCTATCCGGCCACATCA 1249  
QY 241 GGTACCGCATGGTGGGATATGATGATGAATGCTGCTACCTTACAGAGCCCTAAGTGA 300  
DB 1250 GGTACCGCATGGTGGGATATGATGATGAATGCTGCTACCTTACAGAGCCCTAAGTGA 1309  
QY 301 TCGCAGCTACTCCGGATCCCAAGACTGTGTGAGATGAGTGGGGGGCCCACTGGGGA 360  
DB 1310 TCGCAGCTACTCCGGATCCCAAGACTGTGTGAGATGAGTGGGGGGCCCACTGGGGA 1369  
QY 361 GTCCTGGCGGCGCTTGGCTATTTCCATGATGGGGAAGCTGGGCTAAGTCTTGTGTG 420  
DB 1370 GTCCTGGCGGCGCTTGGCTATTTCCATGATGGGGAAGCTGGGCTAAGTCTTGTGTG 1429  
QY 421 ATGCTACTCTTTGGCGGCTTGAACGGGAACTTACACGACAGGGGGGACACACGGCGC 480  
DB 1430 ATGCTACTCTTTGGCGGCTTGAACGGGAACTTACACGACAGGGGGGACACACGGCGC 1489

QY 481 GCCGCCACAGGAGCTTACATCCCTCTTCACTGGGCGGCTCAAGAAATCCAGCTTGTGA 540  
DB 1490 GCCGCCACAGGAGCTTACATCCCTCTTCACTGGGCGGCTCAAGAAATCCAGCTTGTGA 1549  
QY 541 AACACCAACGAGCTGGGACATCAACAGAAATGCTTGAATGCAATGACTCCCTCCA 600  
DB 1550 AACACCAACGAGCTGGGACATCAACAGAAATGCTTGAATGCAATGACTCCCTCCA 1609  
QY 601 ACTGGGTTCTTCCGCGCGCTTCTTACACGACAGAGTTCAATGCGTCCGATGCTCAGAG 660  
DB 1610 ACTGGGTTCTTCCGCGCGCTTCTTACACGACAGAGTTCAATGCGTCCGATGCTCAGAG 1669  
QY 661 GCGATGCGAGCTGCGCGCCCAATTGACAGTTTGATCAGGGGTGGGCTCCATCACTTAT 720  
DB 1670 CCGATGCGAGCTGCGCGCCCAATTGACAGTTTGATCAGGGGTGGGCTCCATCACTTAT 1729  
QY 721 AATGATGCCAGGCTTGGACAGAGAGCCCTTATGTGGCACTAGCACTCAACCGGTGT 780  
DB 1730 AATGATGCCAGGCTTGGACAGAGAGCCCTTATGTGGCACTAGCACTCAACCGGTGT 1789  
QY 781 GGTATCGTCCCGCGTGGAGGTGTGTGCGCCAGTGTACTGTTTCACTTCAAGCCCTGTT 840  
DB 1790 GGTATCGTCCCGCGTGGAGGTGTGTGCGCCAGTGTACTGTTTCACTTCAAGCCCTGTT 1849  
QY 841 GTGTGGGGACAGACCGATGTTTCCGCGCCCTTACGTAAGATGGGCTGAATGAGACG 900  
DB 1850 GTGTGGGGACAGACCGATGTTTCCGCGCCCTTACGTAAGATGGGCTGAATGAGACG 1909  
QY 901 GACGTGCTGCTTCTCAACAACAGCGGCGCCACAGGGGGAACGTGGTGGGCTGTACATGG 960  
DB 1910 GACGTGCTGCTTCTCAACAACAGCGGCGCCACAGGGGGAACGTGGTGGGCTGTACATGG 1969  
QY 961 ATGAATAGCAGCGGTTTCAACAAGAGTGTGGGAGGCCCGCTGTCAACATCGGGGGGCTC 1020  
DB 1970 ATGAATAGCAGCGGTTTCAACAAGAGTGTGGGAGGCCCGCTGTCAACATCGGGGGGCTC 2029  
QY 1021 GGCACACACACTTGTATCTGCCCCACGAGCTGTCCGGAAGCATCCGAGGCCACTTAC 1080  
DB 2030 GGCACACACACTTGTATCTGCCCCACGAGCTGTCCGGAAGCATCCGAGGCCACTTAC 2089  
QY 1081 ACCAATGGGTTGGGGGCTGTTGGTTG 1107  
DB 2090 ACCAATGGGTTGGGGGCTGTTGGTTG 2116

RESULT 5  
LOCUS A32202 2116 bp DNA linear PAT 10-DEC-1996  
DEFINITION NANBH PT polyprotein fragment.  
ACCESSION A32202  
VERSION A32202.1 GI:1926562  
KEYWORDS  
SOURCE Non-A, non-B hepatitis virus  
ORGANISM Non-A, non-B hepatitis virus  
REFERENCE 1 (bases 1 to 2116)  
TITILE Patent: FR 2655990-A 21 21-JUN-1991;  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..2116  
/organism="Non-A, non-B hepatitis virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:12440"  
308..>2116  
/codon\_start=1  
/product="partial NANBH-PT polyprotein"  
/protein\_id="CAA02094.1"  
/db\_xref="GI:1926563"  
/translation="MSTNPKPORKTKRNTNRRPQDVKFPGGQIVGGVYLPRRGRPL  
GVRATKTSRSGRGRQPIPKARQPEGAWAOPQYPMPLVGNBGMWAGMLSPRG  
SRWSGPTPRRRSRNLGKVIDTLTGCPDLMGYIPLVGAPLGAHALAHGVRLD  
GVNVAATGNLPGCSFSIFILALISCLTIPASAYBVRNGIYHTNDCSNLSIYETAD  
MIMHTPGCVPCVRBNSRCVWALPTTLAKDASIPATIRRHVDLLVGAAPCSAMY

VEDLCSVPLVSLFTESPRRQTVODCNCSLYPHVSGHRMAMDMNMSPTALV  
 SOLLRIPQAVDMVAGAHGVLAVLAVSVNMGAKLVLMFLPAGDGPYTGTH  
 GRAHGLTSLFTGPQAKIOLVNTGSMHNTALNDSLOTJGPAALPYRPNAS  
 GCSERMAISCRPIDPQMGPIYVNSHGDQRYCMHVAPOCGIVPALONCGPYC  
 FTSPVAVGTDFGAPYRPMGENETDVLNNTPREPMGFCTMNSTGFTKTCGS  
 PCNIGVGNNTLICPTDFRKHPEATYVYKCGSGPWL"

## ORIGIN

Query Match 100.0%; Score 1107; DB 6; Length 2116;  
 Best Local Similarity 100.0%; Pred. No. 1e-232;  
 Matches 1107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCCTCCCGCTGCTGGTAGCGCTCACTCCACGCTCGCGGCAAGAGCGCAGATCCCC 60  
 1010 TCCTCCCGCTGCTGGTAGCGCTCACTCCACGCTCGCGGCAAGAGCGCAGATCCCC 1069

61 ACTGCACAATACGACGCGCAGCTGATTTCTGTTGGGGCGCTGCTTCTGCTCCGCT 120  
 1070 ACTGCACAATACGACGCGCAGCTGATTTCTGTTGGGGCGCTGCTTCTGCTCCGCT 1129

121 ANGTAGTGGGGATCTCTGGGATCTGTTTCTGCTCTGCTCAGAGCTTCACTCTCG 180  
 1130 ANGTAGTGGGGATCTCTGGGATCTGTTTCTGCTCTGCTCAGAGCTTCACTCTCG 1189

181 CCTCGCGCATGACAGAGCTACAGACTGCAATTTGTTCAATCTATCCCGGCAAGATCA 240  
 1190 CCTCGCGCATGACAGAGCTACAGACTGCAATTTGTTCAATCTATCCCGGCAAGATCA 1249

241 GGTACCGCATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 1250 GGTACCGCATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1309

301 TCGCAGTACTCCGGAATCCCAAGCTGTGTGGAAGATGATGATGATGATGATGATGATGAT 360  
 1310 TCGCAGTACTCCGGAATCCCAAGCTGTGTGGAAGATGATGATGATGATGATGATGATGAT 1369

361 GTCCTGCGGGGCTTGGCTACTATTCATGATGGGGAAGCTGGGCTAAGTCTTGTTGTG 420  
 1370 GTCCTGCGGGGCTTGGCTACTATTCATGATGGGGAAGCTGGGCTAAGTCTTGTTGTG 1429

421 ATGCTACTCTTTCGCGGCTTGAACGGGGAACCTTTACGACAGGGGGGACACAGGCGCG 480  
 1430 ATGCTACTCTTTCGCGGCTTGAACGGGGAACCTTTACGACAGGGGGGACACAGGCGCG 1489

481 GCGGCCACGCGGCTTACATCCCTCTTACACCTGCGGCGCTGAGAAATCCAGCTTGTG 540  
 1490 GCGGCCACGCGGCTTACATCCCTCTTACACCTGCGGCGCTGAGAAATCCAGCTTGTG 1549

541 AACACCAACGCGAGCTGGAATCAACAGAACTGCTTGAATGCAATGATCCCTCCAA 600  
 1550 AACACCAACGCGAGCTGGAATCAACAGAACTGCTTGAATGCAATGATCCCTCCAA 1609

601 ACTGGGTTCTTTCGCGGCTTGTCTACAGCAGAGTTCAATGCGTCCGAGTGTCCAGAG 660  
 1610 ACTGGGTTCTTTCGCGGCTTGTCTACAGCAGAGTTCAATGCGTCCGAGTGTCCAGAG 1669

661 CGCATGGCAGCTGCGGCTTGAACAGTGTGATGAGGGGTGGGGTCCCATATCTAT 720  
 1670 CGCATGGCAGCTGCGGCTTGAACAGTGTGATGAGGGGTGGGGTCCCATATCTAT 1729

721 AATGAGTCCACGCGCTTGAACAGAGGCTTATGCTGGGCACTAGCAGCTCAACGCTGT 780  
 1730 AATGAGTCCACGCGCTTGAACAGAGGCTTATGCTGGGCACTAGCAGCTCAACGCTGT 1789

781 GGTATCGTCCGCGCTTGAAGGTGTGTGCGCCAGTGTACTGTTCACCTCCAGCCCTGT 840  
 1790 GGTATCGTCCGCGCTTGAAGGTGTGTGCGCCAGTGTACTGTTCACCTCCAGCCCTGT 1849

841 GTGTGGGAGACAGCAATCTTTTGGGCGCTTACGTAAGATGGGGTGAAGAAAGAACG 900  
 1850 GTGTGGGAGACAGCAATCTTTTGGGCGCTTACGTAAGATGGGGTGAAGAAAGAACG 1909

901 GAGGTGTGCTTTCACACAGCGCGGCGGCAAGGAGGCAATGCTGCTGCTGATCATG 960

|||||  
 Db 1910 GACGTCTGCTTCTTCAACAACAGCGCGCGCCACAGGGGCAACGTTGCTGTATATG 1969

961 ATGAATAGACCGGGGTTTCAACAAGAGTGTGGGGGCCCCCGGTGCAATCTGGGGGGGTC 1020  
 1970 ATGAATAGACCGGGGTTTCAACAAGAGTGTGGGGGCCCCCGGTGCAATCTGGGGGGGTC 2029

1021 GGCACAACACTTGTGATCTGCCCCACGAGCTGCTCCGGAAGCATCCGAGGCACTTAC 1080  
 2030 GGCACAACACTTGTGATCTGCCCCACGAGCTGCTCCGGAAGCATCCGAGGCACTTAC 2089

1081 ACCAAATGCGGTTGCGGGGCTTGGTTG 1107  
 2090 ACCAAATGCGGTTGCGGGGCTTGGTTG 2116

RESULT 6  
 AR144050  
 LOCUS AR144050 2116 bp DNA linear PAT 08-AUG-2001  
 DEFINITION Sequence 21 from patent US 6210675.  
 ACCESSION AR144050  
 VERSION AR144050.1 GI:15105917  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2116)  
 AUTHORS Hightfield, P. Edmund., Rodgers, B. Colin., Tedder, R. Seton. and  
 TITLE Barbara, J. Anthony, James.  
 JOURNAL Pt-NMB hepatitis B polypeptides  
 Patent: US 6210675-A 21 03-APR-2001;  
 FEATURES  
 source 1..2116  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 1107; DB 6; Length 2116;  
 Best Local Similarity 100.0%; Pred. No. 1e-232;  
 Matches 1107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCCTCCCGCTGCTGGTAGCGCTCACTCCACGCTCGCGGCAAGAGCGCAGATCCCC 60  
 1010 TCCTCCCGCTGCTGGTAGCGCTCACTCCACGCTCGCGGCAAGAGCGCAGATCCCC 1069

61 ACTGCACAATACGACGCGCAGCTGATTTCTGTTGGGGCGCTGCTTCTGCTCCGCT 120  
 1070 ACTGCACAATACGACGCGCAGCTGATTTCTGTTGGGGCGCTGCTTCTGCTCCGCT 1129

121 ATGTAAGTGGGGATCTCTGCGGATCTGTTTCTGCTCTGAGCTTCACTCTCTCG 180  
 1130 ATGTAAGTGGGGATCTCTGCGGATCTGTTTCTGCTCTGAGCTTCACTCTCTCG 1189

181 CCTCGCGCATGACAGAGCTACAGACTGCAATTTGTTCAATCTATCCCGGCAAGATCA 240  
 1190 CCTCGCGCATGACAGAGCTACAGACTGCAATTTGTTCAATCTATCCCGGCAAGATCA 1249

241 GGTACCGCATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 1250 GGTACCGCATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1309

301 TCGCAGTACTCCGGAATCCCAAGCTGTGTGGAAGATGATGATGATGATGATGATGATGAT 360  
 1310 TCGCAGTACTCCGGAATCCCAAGCTGTGTGGAAGATGATGATGATGATGATGATGATGAT 1369

361 GTCCTGCGGGGCTTGGCTACTATTCATGATGGGGAAGCTGGGCTAAGTCTTGTTGTG 420  
 1370 GTCCTGCGGGGCTTGGCTACTATTCATGATGGGGAAGCTGGGCTAAGTCTTGTTGTG 1429

421 ATGCTACTCTTTCGCGGCTTGAACGGGGAACCTTTACGACAGGGGGGACACAGGCGCG 480  
 1430 ATGCTACTCTTTCGCGGCTTGAACGGGGAACCTTTACGACAGGGGGGACACAGGCGCG 1489

QY 481 GCCGCCACGCGGCTTACATCCCTCTTTCACACCTGGCGGCGCTAGAAAATTCAGCTTGTA 540  
 DB 1490 GCCGCCACGCGGCTTACATCCCTCTTTCACACCTGGCGGCGCTAGAAAATTCAGCTTGTA 1549  
 QY 541 AATACCAACGCGAGCTGGACATCAACAGAACTGCTTGAACCTGCAATGACTCCCTCCAA 600  
 DB 1550 AATACCAACGCGAGCTGGACATCAACAGAACTGCTTGAACCTGCAATGACTCCCTCCAA 1609  
 QY 601 ACTGGTCTCTTCCGCGCTGTTCTTACACGACAGGTTCAATGAGTCCGAGCTCAGAG 660  
 DB 1610 ACTGGTCTCTTCCGCGCTGTTCTTACACGACAGGTTCAATGAGTCCGAGCTCAGAG 1669  
 QY 661 CGCATGCGCAGCTGCGCGCCCATTTAGCAAGTTCAGTACGAGGAGGAGGAGTCCCATCTAT 720  
 DB 1670 CGCATGCGCAGCTGCGCGCCCATTTAGCAAGTTCAGTACGAGGAGGAGGAGTCCCATCTAT 1729  
 QY 721 AATGATGCCACGCGCTTGAACCAAGAGCCCTATTGCTGAGCACTAACGACTCAACCGTGT 780  
 DB 1730 AATGATGCCACGCGCTTGAACCAAGAGCCCTATTGCTGAGCACTAACGACTCAACCGTGT 1789  
 QY 781 GGTATCGTCCCGCGCTTGCAGGTGTGTGGCCAGTGTACTGTTTCACTCCAGCCCTGTT 840  
 DB 1790 GGTATCGTCCCGCGCTTGCAGGTGTGTGGCCAGTGTACTGTTTCACTCCAGCCCTGTT 1849  
 QY 841 GTGGTGGGAGCAGCCGATCGTTTCGGCGCCCTACGTAACGATGGGGTGAATAAGACG 900  
 DB 1850 GTGGTGGGAGCAGCCGATCGTTTCGGCGCCCTACGTAACGATGGGGTGAATAAGACG 1909  
 QY 901 GAGCTGCTGTTCTTCAACACACGCGGCGCCCAACGAGGCACTGGTTCGCTGTACATGG 960  
 DB 1910 GAGCTGCTGTTCTTCAACACACGCGGCGCCCAACGAGGCACTGGTTCGCTGTACATGG 1969  
 QY 961 ATGAATAGCACCAGGTTTCAACAGACGTGTGGGAGGCGCCCGGTGCAACATCGGGAGGAGTC 1020  
 DB 1970 ATGAATAGCACCAGGTTTCAACAGACGTGTGGGAGGCGCCCGGTGCAACATCGGGAGGAGTC 2029  
 QY 1021 GGCACACACACTTTGATCTGCCCCACGAGCTGCTTCGGAAGACATCCCGAGGCACTTAC 1080  
 DB 2030 GGCACACACACTTTGATCTGCCCCACGAGCTGCTTCGGAAGACATCCCGAGGCACTTAC 2089  
 QY 1081 ACCAATGCGGTTTCGGGAGCTTGGTTG 1107  
 DB 2090 ACCAATGCGGTTTCGGGAGCTTGGTTG 2116

RESULT 7  
 HPCK1R2  
 LOCUS Hepatitis C virus (strain HCV-1b, clone HCV-R1-R2), complete genome sequence. 9410 bp, RNA linear, VRL 10-FEB-1999  
 DEFINITION  
 ACCESSION D50481  
 VERSION D50481.1 GI:1030705  
 KEYWORDS Polypeptide core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A; NS5B; envelope protein; non-structural protein; interferon-sensitive; interferon-resistant; IFN-sensitive; IFN-resistant; ISDR; interferon sensitivity determining region; HVR; hypervariable region.

SOURCE  
 ORGANISM Hepatitis C virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE  
 AUTHORS Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Izumi, N., Maruno, F. and Sato, C.  
 TITLE Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region  
 JOURNAL J. Clin. Invest. 96 (1), 224-230 (1995)  
 MEDLINE 95340824  
 PUBMED 7542279  
 REFERENCE 2 (bases 1 to 9410)  
 AUTHORS Enomoto, N.  
 JOURNAL Unpublished

# REFERENCE AUTHORS TITLE JOURNAL

3 (bases 1 to 9410)  
 Enomoto, N.  
 Direct Submission  
 Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental University, Second Department of Internal Medicine, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113, Japan (E-mail: EN04522@net.tyase.ort.jp, Tel:03-3813-6111 (ex.3224), Fax:03-3818-7177)  
 Location/Qualifiers

## FEATURES source

1..9410  
 /organism="Hepatitis C virus"  
 /mol\_type="genomic RNA"  
 /strain="HCV-1b"  
 /db\_xref="taxon:11103"  
 /clone="HCV-R1-R2"  
 /note="IFN-resistant HCV from patient 2"  
 330..9362  
 /codon\_start=1  
 /product="polyprotein"  
 /protein\_id="BA09072.1"  
 /db\_xref="GI:1814086"  
 /translation="MSTNPKPQKTKRNTNRBRPDVYKFGGQIVGVVILDPGRPLGVATRKTSERSOPRGRDIPKAPQPRGAAOPEYPLVGNESGMAGMLISRGSRPSGPTDPRRSRLGKVIDTLTTCGRADLNGIPIPLVAPLQASRALARGVIEDGVNATGNLPGCSFSTFLALISCLTTPASAVEVRASGVHTVDSCNASLYEAMDMHTPGCVPCVREANSRCWALPTPLAARNSVPTTIRRHVLDLGAAPCSAMYVGDLCGSVPLVQLFTFSPRHEVTVDCCNSLYPGHSGHRAMMMWMSPTTALVSQLRLPQAVMVAAGAHWGLAGLAAYSGMGMKVLIVMLLFAVDGNTVYSGAKSHHTQGLVSLPAPGQOKTQLVNTNGSMHINLTALNCNLTMLPLAALFTHRRPSCGPERIASCRSIDAPQCGPITTYAERSSDPRPYCMHAPPCGIVPASEVCGCTFSPSPVAVGTDRSGAPVYSGWENETDVLNNTRAPQNMFGCTMMNCTGFTKCGPCPNIGGVANNLTLCPTDCFRKHGREATYTKCSGPMPLPRCAVDYPRLMHPCTVNF TIFKVMYGVGEHRLNAACTNWRGCELEBRDSSELSPLISTEEMILICPSEFTLPLSTGLILHONIVDOVLYGVGSAVVFVKMEYILFLPLADARCAKMMMLL IAOERALENLIVLNAASVAGAHGSPISLPFCMAAYIKGKIVPGAAYFVGMVPLLLILALPPRAYAMDREMAACGCAVPTGALILTLSPRYKYLARLIMWLOYPTTRBAACIQVWLPPLNVRGSDAIILTCALHELIFSTIKLIVALLGLPLMVLAQITKMPFVR AQLIRACMLVRKAGVGYOMAEMKALGTYYVDHLPTQDMVHGLRLDAVAE PVPFSMEETKILTWGADTAAAGDILGLVASRRGKEIFLGADSLGSGWBLAPIT AVSOOTRGLGICITLNGRKNQVGEYQVSTAOPLATCNGSCVTYVHAGASK TLAGBKGPIYQWYTVNDOLVGMQAPSGARSLTPCTCGSSDLIVYTRHADVI PVRRG DGRGSLSPRPVSYLKSGSGPLCPSGHAGVIFRAUVTREYARAVDTPHESMETT MRSVPFTDSSPPAVQTRQVAHLHAPTSGSKNRPVVAQYKAVLVNLSVATL GEGAVMSKAGVDVPIRTGVRTITGAPITVSGYVGFADGGCSGAGVDIILCDECHS TDSSTILGIGVLDQAEITAGARLVLATATPGSVVFPNHLIEVLSMTGEIPFYGK AIPETIKGRLIFCHSKKCDDELAIXSGINAVAYRGIDVSVIPASGQVAVVA TDMATGPTGDRSDVTDCTCYQVYDPSLDPTTETTTVODASVRSORGRGTRV RAGITRFVTTPGERPSGMDPSVLCBCTDAGCAWELTPEITSYRLAAYNTPELPVWQ DHLEWESVFTGLTHIDAFISQTKQAGNPFYLVAVQATCARAQPSPSDQMKC LIRLKEPTLHPPLPLRLGAVONEVILTHPIKTYMAACMSADLEVTSTWVLGVGLA ALAAVCLITGSSVIVIRILISGKPAIIPREVLVYEFDEMEECAHLPIYEGOMLAE OKOKALGLLOTATKOAEAAAVVESKMOALETFPMKHMWNTFSIGTOYLAGISTLBN PALISLMARTASITSPILTQHTLLFNILGAWVAOALAPRAASAPFAGIAGAASISI GLKATVLDILAGTACAGVALVAFKMSGDMSTEBELVNLPAIISFGALVYGVCAAL IIRRHVGEPEGAVQNMNRLIAPASRGNHSPHYVESPAARLVTOILSLTYTQLK RHLQWIMNEDCSPPCSGWLRYDMVLCYLSOPFKWLOSRVLPRLPGVLEFCORRYK GYWRGDIQIOTPCGCAQITGVHKGMSMLIIPREVLVYEFDEMEECAHLPIYEGOMLAE APNYSALIRVAAAEVEYTRVGDPHYVAGMTDNLKCYQVPAPEFELDGAARL YAPASKPELRDVTROYGNOYVVGSLPCEBEPDYNLTSMLTDSHTTARTARRL ARGSPSLTSSASQUSAPSLATCTTHDSDEPADITENLMLRQMGNTITRVESN KVVILDSFPLPAAEDBEREVSIPAEILRSKRFPRAMPAPDPNPPPIESMAGDT VPPVHDSFPLPPTKAPPIPPPRKRITVLTSTESALAEALATKATFSGSESAASQGT ESSIYQCCDLAEARQALISLTERLYTIGPLTNSKQNGYRRCRASGVALTSSCGTLL TCYLKATAACRAKALQDCTMLVCGDVLVYITCSAGQEOBAAALRVTEMTATSPAGP DBPOPEYDELITSSNSVVAHDHAGVLRVYLTDPPTPLARAETRAHPFVNSML GNITNAPTILAKRMILMTHPSIILAQBLEKALDCQITGACYSIEPLDLPQIQGLH GLSAPSLHSYSGEINRVASCIKRGVPLLRWRHARSVRAKLISGGRAALCGKTL FNMVAVRTKLTLPPIPAASQDLSGWFVAGYSGDIDYHSVSRARPRFMWCLLLISGV GYLLPNR"





GVNATGNTLPGCSFSLFLLALLSLCTLPASALEVRNVSGVHTVNDCSNAIIVEYEAAD  
MIMHTPGCVPEVRNENSSRCVATLPETLAARNSSIPPTTIRRVDDLMMVSPALFCSMY  
VGLDGSFVLSQLFTEPSRREHTVODCNCSIRPGHSGVMGMMMAITLMTALVAV  
SOLLRIRPOAVBNVAGAHMGVLAGLAYSAGNAGMAKVLIMLIFAVDDEBTYTTGGAO  
AHTTRGLATFSFGRSONOQLINTNGSMHINRTALCONDSINLNGFLAALPYARFNAS  
GGERNASCRPIDKFAQMGPIATYAKPLSDQKPYCMHTAPQPCITVPASQVCSPYIC  
FPSPVAVGTDDFVFPYTSWGENETDVLINNTPRPKMGKFCCTMMNNGKFPKTCGG  
PDCDIDGAVGNDTLICPTDCFRKHEATYVCGSGPMITRRCMDVYRGLMHYCTANF  
TIPKMYVYGVEHRLNLAACNMTGRGCEDEDRSELBPILSTTEMOWLPCSPFTLL  
PALSTGLIHLHRIHVVOYLYIGTGSAAVSAIKWEYVLLFLLEALRACACMMMLL  
IYQAEALENLVILNAAVAGTHGIIISPIVFCAAATYIGRLVPGAAAYAPXGWPLL  
LLIAPLPRAVAMREMASSCGGAVFTGLALLTSPHYKPLFALIMWLQFTIRARAH  
LQVMIPLNVRGGDAIILMCAIHEBLFATIKLILGLIMLVLOAGITRVPYVR  
AHGLIRACVAVREVAGHYVOMALMLAALTGYVDHLITPDMHEGRLDAIYVE  
PVESSMETKLIITWAGDTAACGDIILGLPVSARGREILILGDSGOMRLAATIT  
AYAOGRGLIGCTVSLTGRDNQVGEVNUVSTASPLATCVNVCMTVYHGAQTK  
TLAAGQPTTQNTTNDODLVCHLAPPGARSLPTCTGSSDLITVRHADVIFVRERG  
DSRGSLLSPRPVSYLKGSSGGPLLCPSGHAUVGI.FRAVACSTRGAKVAVLVPVPSMETT  
MRSPFTDNSSPPAVPQTFQVAHMAPTGSGSTKVPAAVAAQGVAVLVPVPSMETT  
GFGAAMSKAHVDPNIRTGVRTTTGAPITVSTYKELADGSGGSAVDIILDEGHS  
TOSTTLLIGTVIDOAEHTAGARLVLATATPGSVTPVPHNIEEVALNSINIGELPYGK  
AIIPIETIKGRHLIFCHSRKCKDEBLAAKLSGLMANVAYTRGIDVSVITPSGUVVVA  
TDLMTGTGDEPDSVIDCNTCVIQTVDPSLDPTFTIETITVPDPAVSQGRRTGRG  
RTGIIYRFPVTPGERSGMFPDSVLCCECYDAGCAWYELTPAETSURLAVALNTPGLPVQ  
DLLEFWEGVFTGLHIDAHLSQTKQAGNFPYLVAVOATVCAARAQAPSPMDQMKC  
LILKPLTNGPPTLRYLAGVONEVILHPIITKFIAMCMADELVATSTVILVGVILA  
ALAAVCLTTSVVIVGRILISGKEPAIVPDEVLVYOPDEMEBEACHTLPIYBOGMOAB  
QFOKALGLIQTATKQAEAAAPVESKWRALSEFPAKHMNFIISGIOYLAGSLTILPGN  
PATIASLMAFTASTPLTTOHTLIFNLGGMVAQALPSPAASAFAGAGIAGVAVCSI  
GLGKVLVLDIAGGAGVAGALVAFKVMGSGMSTEDLVNLPILSPBAGLVAGVAVCA  
ILRHVYPGEGAVOMMNRILAFASRGHVSPTHVPSDIAVAVTQIISLITTOILK  
RLHOWINDECSPTGSGMLDWDMICTVLTIDPMSQSLPRLGVPPLSGORXK  
GVWRGGIMOTTCPCGAOITGVHNGSMRIVGRTCSNTHMGHPINAAITVGTCTSP  
APVYSALMRVTABEYILEVTOVEDFHYVVTGTTDNKCPQVAPBEPFTVDSVRLHR  
YAPCPKPLEBEVTFVGLNLYVGSOLPCEPBDVAVLTSMLIDSHITAEAKREL  
ARSPSPILSSASSQLSAPSLKATCTTRRSGPADLIEANLIRMGEMGATITVESEN  
KXVILDSFEBLRAEEDEREVSAPAILRBSRKPRAPIIARDVNPPLVESMGRT  
VPPVHVGCLPPTKAPPIPPRRKRTVLTSTVSSALATKTSGSSSDSVSDGCT  
ATAPOPSDDPDGAGSDVESGSMPTVSEGRDPLSDSGSVTSEASDVSVCMS  
YTTGALITPSCSAEBSKLPINALSNSILRHNNVYATITSSAQRQKKTFRDLQVLD  
DHYDVLKEMKAASTVAKALPVEBECKLTPHSAKSKFYCAKADKAVRNISCAVNI  
LSVWKDLEBTEPITDTTMAKNEVCEVQDEKGRKPARLIVPDLGVRNCEMAALYD  
VVSITLPOAVMSSSYEQSPQORVEFLVPLNKKKKKPMGAVATRCDSVTSIDRLV  
EESIYOCCLAPBAROARISLTERLYGGLTNSKQNGCYRGRASGVLTTCGNTL  
TCYLKSAACRAKLODCMTLNGCDLVLVCEASGTTEPASLRVTEAMTARTVSNML  
DPRKPEYDELTSCSSNVSVAHDASGRVYITTRDPTTFLAAMATEARTVSNML  
GNLIMVAPTLIMAMILLMTHFSLILAQOEKALDCQIYAGCSISEPLDLPQIIRLH  
GLSAFSLHSYSPGEINRVACLRKLGVPLRVRHRAVSRAKILISQSGRANACGYL  
FNMAVTRKTLKLTPIPAASQDLSSWFAVAGSGDIYHLSLRABRPMFLCLILSYGV  
GIYLLPNR"

ORIGIN

Query Match	85.5%; Score 946; DB 14; Length 9359;
Best Local Similarity	90.9%; Pred. No. 2e-197;
Matches 1006; Conservative	0; Mismatches 101; Indels 0; Gaps 0;
QY	1 TCTCCCGGCTGCTGGGTAGGCTCACTCCCAAGCTCGCGGCAAGAGCGCAGATATCCC 60
DB	1032 TCTCCCGGCTGCTGGGTAGGCTCACTCCCAAGCTCGCGGCAAGAACTCCAGATATCCC 1091
QY	61 ACTGCAATATAGAGCAAGCAAGTGAATTGCTGTTGGGGGGGCTGCTTCTGCTCGCT 120
DB	1092 ACTACGCAATATAGCACTGACGATTTGCTGCTTGGGAGCGCTGCTTCTGCTCGCT 1151
QY	121 ATGTACGTGGGGAGTCTCTGCGATCTGTTTCTCTGCTCTCAAGCTTTCACTTCTCG 180
DB	1152 ATGTACGTGGGAGATCTMTGGAGTCTGTTTCTCTGCTCTCAAGCTTTCACTTCTCG 1211
QY	181 CTTGCGGCAATCAAGACGCTACAGACTGCAATTTGTCATCTATCCCGGCAAGATCA 240
DB	1212 CTTGCGGCAATGAGACAGTACAGACTGCAATTTGTCATCTATCCCGGCAAGATCA 1271
QY	241 GGTACCGCAATGGCTTGGGATATATATGAACCTGCTCACTCAAGAGCCCTAAGTGTGA 300

DB	1272 GGTACCGCATGCTTGGGACATGATGAAGAACTGCTCACTAACAACACTTAACTGATG 1331
QY	301 TCGAGCTACTCCGAGTATCCCAAGCTGCTGAGACATGATGAGCGGGGCCCACTGGAGA 360
DB	1332 TCGAGTACTCTCGGATCCCAAGCTGCTGAGACATGATGAGCGGGGCCCACTGGAGA 1391
QY	361 GTCCTGGCGGGCTTGCCTACTATTCATGATGAGGGGAACTGAGGCTTAAGGCTTGTGTG 420
DB	1392 GTCCTGGCGGGCTCGCTACTATTCTATGAGCGGGGAACTGAGGCTTAAGGCTTGTGTG 1451
QY	421 ATGCTACTCTTGGCGGGCTTGAACGGGGAACCTTACAGCAAGGGGGGACACAGCGCCG 480
DB	1452 ATGCTACTTTTGGCGGGCTTCAACGGGGAACCTTATACACAGGGGGGGGAGCGCCAT 1511
QY	481 GCGGCGCAAGGGCTTACATCCCTTCAACACCTGGGGCGGCTCAGAAATCCAGCTTGA 540
DB	1512 ACTACCGGGGGCTTGGCAACCTCTTTTCTGTTGGCGGCTCCAAACATCAAGCTTATA 1571
QY	541 AACACCAAGGAGCTGAGCATCAACAGAACTGCTTGAACCTGCAATGACTTCCCTCAA 600
DB	1572 AACACCAAGGAGCTGAGCATCAACAGAACTGCTTGAACCTGCAATGACTTCCCTCAAC 1631
QY	601 ACTGGCTTCTTGGCGGCTGCTTCTACAGCAAGCTTCAATGCTTCCGATGCTCAAG 660
DB	1632 ACTGGCTTCTTGGCGGCTGCTTCTACAGCCCAAGATTCAACGCGTCCGATGCGGAG 1691
QY	661 CGGATGGCAAGTGGCGGCGCCCAATGACAGTTCGANTCAGGGGAGGGGTCCTATCACTTAT 720
DB	1692 CGGATGGCAAGTGGCGGCGCCCAATGACAGTTCGANTCAGGGGAGGGGTCCTATCACTTAC 1751
QY	721 AATGATGCCAAGGCTTGGACCAAGAGGCCCTTATTTGCTGCACTACGCACTCAACCGTGT 780
DB	1752 GCTAACCCGCTCAGCTTGGACCAAGAGGCCCTTATTTGCTGCACTACGCACTCAACCGTGC 1811
QY	781 GGTATGTCGCCGCGCTTGCAGAGGTGTGTGGCCCAAGTGAATCTGTTTCACTCAAGCCTGTT 840
DB	1812 GGTATGTCGCCGCGCTGCGAGGTGTGTGGCTTCAAGTGAATCTGTTTCACTCAAGCCTGTT 1871
QY	841 GGTGTGGGAGCAACGCAATGCTTTCGCGGCGCCCTAAGTCAAGATGGGGTGAATGAGAG 900
DB	1872 GGTGTGGGAGCAACGCAATGCTTTCGCGGCTTCAAGTGAATGGGGGAGAAATGAGAG 1931
QY	901 GACGTCTCTCTTCAACAACAGCGCGGCCCAAGCGGGCAACTGTTCCGCTGTATCATG 960
DB	1932 GACGTCTCTCTTCAACAACAGCGCGGCCCAAGCGGCAACTGTTCCGCTGTATCATG 1991
QY	961 ATGAATAGCAACCGGCTTCAACAAGAGCTGTGGGGGCGCCCGCTGCAACATCGGGGGGCTC 1020
DB	1992 ATGAATAGCAACCGGCTTCAACAAGAGCTGTGGGGGCGCCCGCTGCAACATCGGGGGGCTC 2051
QY	1021 GGGCAACAACCTTGAATCTGCCCCCAAGCACTGTTCCGGAAGCATCCGAGCACTTAC 1080
DB	2052 GGGCAACAACCTTGAATCTGCCCCCAAGCACTGTTCCGGAAGCACTCCGAGCACTTAC 2111
QY	1081 ACCAAATGCGGTTGCGGGGCTTGGTTG 1107
DB	2112 ACCAAATGCGGTTGCGGGGCTTGGTTG 2138
RESULT 9	
HPCK1S2	9410 bp RNA linear VRL 10-FEB-1999
LOCUS	Hepatitis C virus (strain HCV-1b, clone HCV-K1-S2), complete genome
DEFINITION	sequence.
ACCESSION	D50485.1 GI:1030704
VERSION	
KEYWORDS	polyprotein; core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A; NS5B; envelope protein; non-structural protein; interferon-sensitive; interferon-resistant; IFN-sensitive; IFN-resistant; ISDR; interferon sensitivity determining region; HVR; hypervariable region.
SOURCE	Hepatitis C virus



ORGANISM Hepatitis C virus  
REFERENCE 1 (bases)  
AUTHORS Yamamoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Izumi, N., Marumo, F. and Sato, C.  
TITLE Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region  
JOURNAL J. Clin. Invest. 96 (1), 224-230 (1995)  
MEDLINE 95340824  
PUBMED 7542279  
REFERENCE 2 (bases 1 to 9410)  
AUTHORS Enomoto, N.  
JOURNAL Unpublished  
REVIEWER 3 (bases 1 to 9410)  
AUTHORS Enomoto, N.  
TITLE Direct Submission  
JOURNAL Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental University, Second Department of Internal Medicine, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113 (ex.3224), Japan (E-mail:PXN04522@iitfyserv.eor.jp, Tel:03-3813-6111 (ex.3224), Fax:03-3818-7177)  
FEATURES Location/Qualifiers  
1. .9410  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/strain="HCV-1b"  
/db\_xref="taxon:11103"  
/clone="HCV-K1-S2"  
/note="IFN-sensitive HCV from patient 2"  
330. .9362  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="BA09076.1"  
/db\_xref="GI:1814089"  
/translation="MSTNPKPQRKTKRNTNRPDVYKFGGQIVGGVILLPRGRPLGVATRTKTSRSPQRRGROPKPARQPEGRAMQPGPMPLYNEGAGMGLSFRGSRWGPPTDRRRRNMGKVIDTLTGCFALMGIYPLVGAFLGASRALAHGVRLDGVNATGNLPGCSFSTFILALSLCTIPASAYEVRAAGVHYHTNCSNASTYVEADMIMHTGCPVCPREANSRCWVALTPTLAARNSVPTTTRRVLDLVGAACSAHYVGLDGSFVLSQLFTFSPRHETVDDNCSTIPGVSHGRMAMMMNSPTTLAVSOLRLPOAVMDVAGAHGVLAVLAIVSVGWMAKYLVMFLFAGVDGTTVTGAQGHRTORLASLFTGAQOQRIQLVNTGSMHINPALCNBSLNTGFLALFYTRFNSGCEPERJASCSIDAPAGMGPIYAEBSGSDORPCMHYAPRGGIYVASEVGPYCFTPSPVVGTGTRSGAPDYSGENETDVLINNTBAPQGMPCCTMMNGCFKTKGCPENICGVGNNTTCTPTDCPRKPEATYTKCSGGPDTPECMVDYTRVNHPTCTNF1LEKRMVYGVGEHRLAACNMTGRSCLEDRELSLSTTEMQILPSCFTFLPVLSGLIHQNIIDVOYLVGVSAVSVFVMEYILFLFLADARVACAMMLIAQAEALENLVINASVAGAGLFLVPCAWYIKGLVPGAAVAYGVPMILLLALPRAVYAMDRMAASCGGAVFVGLALITLSPHYKVLALIMWLOVETIRACLOVWIRPLAVRGRDAIILITCALHPELISITIKIILVAILGPMVIOAGITKPYRVAQGLRACMLVRKVGAGHYVQAFKFLALITGVVDHLTPLODMAMHAGRLDAVEPVFSDMETKIIITWADTAACGDIILGLAVSARGKEIFLGPDSLEGSGWRLAIPITAVSOQTRGLGCIITSLTGDKNOVEGVQVSTASFLATGPNLCMVTVVGAISKTLAAGSGPIITOMYTNVDDLVGOWAPSGARSLPCTCGSSDILVTRHADVIRPVRGDRGSLSPRPVSYLKSGSGGLCPSGAGVGIERRAICTRGAVKAVDEIPVSMETTMESEPVTNDSPPRAVPTQVYALHAPITSGSKNKPIPERVAAQGYKVLVINEVATLGGFAYSKAHGVDNPIRTGRTITTGAPITYSYTGKFLVDGSGGAGVNDIICDECHSTOSTILGIVYDQAEIAGARLVLAITPISGVLPADNIEVALSINGELPDECHSAIPERTIKGGRHLIFCHSKKCCDELAJSLGILINAAVYRGDVAVIPASGVVVA TALMTGTFGDFSVIDCNTCVQIVDFSLDPFTTETTTVPODARSORGRGTGRAGIYRFTVPRGSRGMPDSVYCECYDAGCAWELTPESTSRVLAIPVWOC DHLEPESVFTGTGTHDAHFLSOTKQAGNFPILVAYQATVCAKAAQAPPSNDMMKCLRLKPLHGPFTFLYRLGAVQNEVILHETIKYIMACSADELEVSTWLVGVIALAAAYCLTGGSVIVGRILLSGRKAIIPREVLRYREDEMECHSAIPYIEMOGLAE OFOKALGLQIOTKQAEAAAPVESKAIPLFTFAKHMNFISGLOIYALSTLJEN PAIASLMAFTASITSPITQHTLLFNILGWMVAOLAPRAAIPFAGAGIAGVACSI GAKVLVDIAGAGVAGVAVAFKVMGSMPESTEDLVNLPAILSGALVGVVAAI ILRHHYBSRGANQVMNRILAPASRGNHVSPTHYVPESDAAAVTOISLITTOILK RLHMOINEDSTPCSGSWLRDWMCTVLTDRKTLVQSLVRLFGVPLTSGCQKPG GWVRGDIQOTTCPGAQITGVHNGSMRVGRTVPCNTSMTHGFLINAYITLPGRTLR APNYSRLMRVAEVEVYRVGDFHYVGMETDNLKPCQVAPPEFKLDEKRLRLR VPASRPLRLDEVTFQVGLNQVYVGSQLPCEPBDVAVLITSLMLTDSHITAAEAKRLL

ARSGPSSLASSASOLASPSLKAATCTTRHSDPADLIEANLIMROEMGNITRESEN KVIILDSFELRLAEDEBEREVSIAEILIRKSRKPPMPPIWARDYNDPILSEKWDY VPPVILDSFELPPTKAPPIPPRKRKTIVLIESVSSALALATKTRGSSSSADSGT ATAAPDOPNDGDGSDVESYSMPLDEGRPDIDSDGSWTVSEAGDVCMS YTTTGALLTPCAEESKLPINALSNLIRHNNVYITTSASQROKQTFPDLQVTD DHRDVLKEMKAASTVKAHLVSBEACKTPHSAKSKFGYAGAKDRNLSRAINI RSWKOLDEBTEPIDITTIMAKNEVCVPEKGRAPRILVPEDLGRVCEKRALD VSTLPOAWMPSEYFGSPGQREVLVNAKSKCPMGFADPTRCDSVTESDIRV EBSIYOCCLAPAROAIRSLRILYIGPGLTNSKCONCGYRCRAGSVLTSCGNTL TCYIKAPACRAKIODCTMLNCGDPLVYICSAQGOEDASIRVTEAMTRYAPPG DPQPEYDLELITSCSNVSAHDAAGKRYIYITRPTPLAAMETARHTVNSML GNIMYAPTLMAEMILMTFSLILAOEULEKLDQIYGCYSIEPLDLPQILQRLH GLSAFSLHSYSGEINRVASCLKLGVPPLRVRRHARSVRATLISOGGPAALCGKYL FNVAVRTKLKLTPIPAASQLDLSGMFVAGSGDIYHVSARAPRPMWMLLISYGV GYLHPNR"  
330. .902  
/product="core protein"  
mat\_peptide 903. .1478  
/product="E1"  
mat\_peptide 1479. .2756  
/product="E2"  
mat\_peptide 2757. .3407  
/product="NS2"  
mat\_peptide 3408. .5300  
/product="NS3"  
mat\_peptide 5301. .5462  
/product="NS4A"  
mat\_peptide 5463. .6245  
/product="NS4B"  
mat\_peptide 6246. .7586  
/product="NS5A"  
mat\_peptide 7587. .9359  
/product="NS5B"  
misc\_feature 1479. .1559  
/note="hypervariable region (HVR)"  
misc\_feature 6954. .7073  
/note="interferon sensitivity determining region (ISDR)"

ORIGIN

Query Match 85.0%; Score 940.6; DB 14; Length 9410;  
Best Local Similarity 90.6%; Pred. No. 3.1e-196;  
Matches 1003; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

1 TCCTCCGCGTGGGATGAGGCTCACTCCACGCTGGCGGCAAGGACGCACATCCCC 60  
1032 TCCTCCGCGTGGGATGAGGCTCACTCCACGCTGGCGGCAAGGATCTCACGCTCCA 1091  
61 ACTGCACAATACGACGCGCACGTCGATTTGCTGTTGGGCGGCTCCTTCTGCTCGCT 120  
1092 ACTACGACAATACGACGCGCACGTCGATTTGCTGTTGGGCGGCTCCTTCTGCTCGCT 1151  
121 ATGTACGTGGGGATCTCTGCGGATCTGTTTCTCTGCTCTCACTGTTACCTTCTCG 180  
1152 ATGTACGTGGGGATCTCTGCGGATCTGTTTCTCTGCTCTCACTGTTACCTTCTCG 1211  
181 CCTCGCGCATACGACGCGTACGACGTCGAATGTTCATATCCCGGCGCATATCA 240  
1212 CCTCGCGCATACGACGCGTACGACGTCGAATGTTCATATCCCGGCGCATATCA 1271  
241 GGTACCGCATGCTGGGATATGATGATGAATGCTGTCACCTTACAGACGCTTACTGTA 300  
1272 GGTACCGCATGCTGGGATATGATGATGAATGCTGTCACCTTACAGACGCTTACTGTA 1331  
301 TCGACACTACTCCGATCCCAACGCTGCTGTGACATGAGTGGCGGGGCGCCACTG33GA 360  
1332 TCGACACTACTCCGATCCCAACGCTGCTGTGACATGAGTGGCGGGGCGCCACTG33GA 1391  
361 GTCTCGCGGCGCTTCCGCTTATTCATAGTGGGGAACCTGGGCTAAGGCTTGGTGG 420  
1392 GTCTCGCGGCGCTTCCGCTTATTCATAGTGGGGAACCTGGGCTAAGGCTTGGTGG 1451  
421 ATGCTACTTTCGCGGCTTGAAGGGAACCTTACACGACGAGGGGGAACACGCGCGCG 480

```

Db      1452 ATGCTACTTTTGGCCGCGCTTGATGGGCGCACCAACCGTGAACAGAGGGGGCGACAGGCCAT 1511
QY      481 GCGGCCACGGGGCTTACATCCCTCTTCAACACTGGGCGGGCTCAGAAATCCAGCTTGTGA 540
Db      1512 ACCACCCAAAGGCTCGCGTCTCTTTACATTTGGGGCGGACAGAGATCCAGCTTGTG 1571
QY      541 AACACCAACGGGAGCTGGACATCAACAGAACTGCTTGAATGCAATGACTCCCTCCA 600
Db      1572 AATACCAACGGGAGCTGGACATCAACAGAGACTGCCGTGAATGCAATGACTCCCTCAAC 1631
QY      601 ACTGGGTTCTTCCCGCGCTGTTCTACAGCACAAGTTCAATGCTCCGATGCTCAGAG 660
Db      1632 ACTGGGTTCTTCCCGCGCTGTTCTACACACAGGTTCAACTGTCGGATGCCACAG 1691
QY      661 CGCATGGCCAGCTGGCGGCCCATTTGACCAAGTGGATCAGAGGGTGGGGTCCCATCACTTAT 720
Db      1692 CGCATGGCCAGCTGGCGGCTTCATCAGACGCTTGGCTCAGGGGTGGGGTCCCATCACTTAC 1751
QY      721 AATGAGTCCACAGGCTTGGACACAGAGGCCCTATTGTGTCGACATACGACCTCAACCGTGT 780
Db      1752 GCTGAGCGGGGAGCTCGGACACAGAGGCCCTATTGTTGGACATACGACCTCGACCGTGT 1811
QY      781 GGTATGTCGTCGCGCTTGGACAGTGTGTGCGCCAGTGTACTGTTCACTCCAGCCCTGTT 840
Db      1812 GGTATGTCATCCCGCGCTCGAGGTGTGTGTGTCAGTGTATTGCTTCAACCCCAAGCCCGT 1871
QY      841 GTGGTGGGGAGCAGACCATGCTTTCGCGGCCCTTACGTACAGATGGGGTGAATGAGAGC 900
Db      1872 GTGGTGGGGAGCAGACCATGCTTTCGCGGCCCTTACGTATAGCTGGGGAGGAATGAGAGC 1931
QY      901 GACGTGCTGCTTTCACACACACAGCGCGCCGACAGGGGCAACTGTTCGGCTGTACATG 960
Db      1932 GAGTGTGTGTCCTCAACACACAGCGGGGCGCCGAAAGGCAACTGTTCGGCTGTACATG 1991
QY      961 ATGATATGACACCGGGTTCACCAAGACGTGTGGGGGCCCCCGCTGCAACATCGGGGGGT 1020
Db      1992 ATGATATGACACCGGGTTCACCAAGACGTGTGGGGGCCCCCGCTGCAACATCGGGGGGT 2051
QY      1021 GGCACACACACTTGTATCTGCCCCAGGAGTGTTCGGAAGACATCCCGAGGCACTTAC 1080
Db      2052 GGCACACACACTTGTATCTGCCCCAGGAGTGTTCGGAAGACATCCCGAGGCACTTAC 2111
QY      1081 ACCAAATGCGGTTTCGGGGCTTGGTTG 1107
Db      2112 ACAAAATGCGGTTTCGGGGCTTGGTTG 2138

RESULT 10
AY746693      1280 bp      mRNA      linear      VRL 10-OCT-2004
LOCUS      Hepatitis C virus clone #09-6 polypotein mRNA, partial cde.
DEFINITION      AY746693
ACCESSION      AY746693
VERSION      AY746693.1 GI:53801768
KEYWORDS
SOURCE      Hepatitis C virus
ORGANISM      Hepatitis C virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Hepacivirus.
REFERENCE      1 (bases 1 to 1280)
AUTHORS      Chambers,T.J., Fan,X., Droll,D.A., Hembrador,E., Slater,T.,
            Nickells,M.W., Duestin,L.B. and DiBisceglie,A.M.
TITLE      E1/E2 quasispecies heterogeneity as a pretreatment variable during
            pegylated interferon therapy of chronic HCV infection
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 1280)
AUTHORS      Chambers,T.J., Fan,X., Droll,D.A., Hembrador,E., Slater,T.,
            Nickells,M.W., Duestin,L.B. and DiBisceglie,A.M.
TITLE      Direct Submission
JOURNAL      Submitted (10-SEP-2004) Molecular Microbiology and Immunology,
            Saint Louis University School of Medicine, 1402 South Grand Avenue,
            St. Louis, MO 63104, USA
FEATURES
            source      1..1280

```

## ORIGIN

```

Query Match      84.4%; Score 934.2; DB 14; Length 1280;
Best Local Similarity 90.2%; Pred. No. 1e-194;
Matches 999; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```

## CDS

```

/organism="Hepatitis C virus"
/mol_type="mRNA"
/db_xref="taxon:11103"
/clone="#09-6"
/country="USA"
/notes="genotype: 1b"
<1..>1280
/notes="contains envelope proteins E1 and E2"
/codon_start=1
/product="polypeptide"
/protein_id="AAU94057.1"
/db_xref="GI:53801769"
/translation="YEVNAGCAHYTNDGNSIYETMDMHTPGVPCVREGNS
SRGWALPTTLAARNASIPYTIIRRHVDLLVAAAFCSVMYVDLCGVFVLSQFTF
SRBRHVDDNCISYIPGHSRHMAMMMMSPTTLAVSOLRIPOAVLVDSQ
HGVVAGLAAYSMVGMAVLIVMLFAGVDGHTTGGSAARNAYGLTTFTRPSQ
KIOLVNTGSMHINRTALNDSLDTGTFALFYHSPASGCPERMAISCPRIENAO
GMGPIITYPEPHSDRPEYKHVAPQPCGTTIPAAQVCGPYCTPSPVYVGTDRCAP
TYNMANETVDVLIINNTPPQGNMFGCTMNGTGTCTCGGFPCKIGVGNNTLICPT
DFFRKHPEATYTKGSGPWLTPRCIVDPYRLMH"

```

```

QY      1 TCCCTCCCGCTGGCTGAGTACGCTCACTCCCAAGCTGCGGCAAGAGCCAGCATCCCC 60
Db      130 TCCCTCCCGCTGGCTGAGTACGCTCACTCCCAAGCTGCGGCAAGAGCCAGCATCCCC 189
QY      61 ACTGACAAATACAGACGCAAGTGTGCTGTGGGGCGGCTGCTTGTGCTCGCT 120
Db      190 ACTGACAAATACAGAGTCAAGTGTGCTGTGGGGCGGCTGCTTGTGCTCGCT 249
QY      121 ATGTAGTGGGGGATCTCGCGGATCTGTTTCTCTGCTCTCAGGTTCACCTCTG 180
Db      250 ATGTAGTGGGGGATCTCGCGGATCTGTTTCTCTGCTCTCAGGTTCACCTCTG 309
QY      181 CCTCGCGCAATCAGACGCTGACAGACTCAATTGTTCAATCATCCCGGCACTATCA 240
Db      310 CCTCGCGGCAATCAGACGCTGACAGACTCAATTGTTCAATCATCCCGGCACTATCA 369
QY      241 GGTACCGGATGGCTTGGATATGATGAACTGTCACTTACACAGACGCTTATGTTGA 300
Db      370 GGTACCGGATGGCTTGGATATGATGAACTGTCACTTACACAGACGCTTATGTTGA 429
QY      301 TCGAGCTACTCGGATATCCCAAGCTGTGAGAAATGTTGGCGGGGGCCCACTGGGA 360
Db      430 TCGAGCTACTCGGATATCCCAAGCTGTGAGAAATGTTGGCGGGGGCCCACTGGGA 489
QY      361 GTCCTGAGCGGCTTGTCTACTATTCATGATGGGGAAGCTGAGGCTTGTGTTG 420
Db      490 GTCCTGAGCGGCTTGTCTACTATTCATGATGGGGAAGCTGAGGCTTGTGTTG 549
QY      421 ATGCTACTTTTGGCCGCGCTTGAACGGGGAACCTTACACGACAGAGGGGGACACAGCCGC 480
Db      550 ATGCTACTTTTGGCCGCGCTTGAACGGGGAACCTTACACGACAGAGGGGGACACAGCCGC 609
QY      481 GCGGCCAAGGGGCTTACATCCCTCTTCAACACTGGGCGGCTGTAGAAATCCAGCTTGA 540
Db      610 AAGGCTTACGGGCTGACAGACCTCTTCACTCGGGGCGGCTGTAGAAATCCAGCTTGA 669
QY      541 AACACCAACGGGAGCTGGACATCAACAGAACTGCTTGAATGCAATGACTCCCTCCA 600
Db      670 AACACCAACGGGAGCTGGACATCAACAGAACTGCTTGAATGCAATGACTCCCTCCA 729
QY      601 ACTGGGTTCTTCCCGCGCTGTTCTACAGCACAAGTTCATGCTCCGATGCTCAGAG 660
Db      730 ACTGGGTTCTTCCCGCGCTGTTCTADACACATAGCTTAAACCGCTCCGATGCCAGAA 789
QY      661 CGCATGGCCAGCTGGCGGCCCATTTGACCAAGTGGATCAGAGGGTGGGGTCCCATCACTTAT 720
Db      790 CGCATGGCCAGCTGGCGGCCCATTTGATGATGCTCAGGGGTGGGGTCCCATCACTTAC 849

```

QY 721 AATGAGTCCACGGCTTGGACAGAGGCCCTAATGTGCTGACCTAAGCAGCCTCAACCGTGT 780  
 DB 850 ACTGAGGCCCAACGGCTTGGACAGAGGCCCTAATGTGCTGACCTAAGCAGCCTCAACCGTGT 909  
 QY 781 GGTATCGTGGCCCGCTTGGACAGTGTGTGGCCAGTGTACTGTTCACTTCCAGGCCCTGT 840  
 DB 910 GGTATCGTGGCCCGCTTGGACAGTGTGTGGCCAGTGTACTGTTCACTTCCAGGCCCTGT 969  
 QY 841 GTGTGTGGGACGACCGATCGTTTGGCCCTTACGTACAGATGGGGTGAATGAGACG 900  
 DB 970 GTGTGTGGGACGACCGATCGTTTGGCCCTTACGTACAGATGGGGTGAATGAGACG 1029  
 QY 901 GACGTGCTGCTTCAACACAGCCGCGCCACAGGGGCAACTGTGCTGCGGTGACATG 960  
 DB 1030 GACGTGCTGCTTCAACACAGCCGCGCCACAGGGCAACTGTGCTGCGGTGACATG 1089  
 QY 961 ATGATATAGCACCGGGTTTCAACAGACGTGTGTGGGGCCCGCTGTCACATCGGGGGGCT 1020  
 DB 1090 ATGATATAGCACCGGGTTTCAACAGACGTGTGTGGGGCCCGCTGTCACATCGGGGGGCT 1149  
 QY 1021 GGGCAACACCTTGTATGTGCCCCAGGACTGCTTCCGGAAGCATCCCGAGGCACTTAC 1080  
 DB 1150 GGGCAACACCTTGTATGTGCCCCAGGACTGCTTCCGGAAGCATCCCGAGGCACTTAC 1209  
 QY 1081 ACCAATGGGGTGGGGGCGCTTGGTTG 1107  
 DB 1210 ACCAATGGGGTGGGGGCGCTTGGTTG 1236

RESULT 11  
 LOCUS AY746700 1280 bp mRNA linear VRL 10-OCT-2004  
 DEFINITION Hepatitis C virus clone #09-40 polyprotein mRNA, partial cds.  
 ACCESSION AY746700  
 VERSION AY746700.1 GI:53801782  
 KEYWORDS  
 SOURCE Hepatitis C virus  
 ORGANISM Hepatitis C virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.

REFERENCE 1 (bases 1 to 1280)  
 AUTHORS Chambers, T.J., Fan, X., Droll, D.A., Hembrador, E., Slater, T.,  
 Nickell, M.W., Duetin, L.B. and DiBisceglie, A.M.  
 E1/E2 quasispecies heterogeneity as a pretreatment variable during  
 pegylated interferon therapy of chronic HCV infection  
 Unpublished  
 JOURNAL 2 (bases 1 to 1280)  
 AUTHORS Chambers, T.J., Fan, X., Droll, D.A., Hembrador, E., Slater, T.,  
 Nickell, M.W., Duetin, L.B. and DiBisceglie, A.M.  
 Direct Submission  
 Submitted (10-SEP-2004) Molecular Microbiology and Immunology,  
 Saint Louis University School of Medicine, 1402 South Grand Avenue,  
 St. Louis, MO 63104, USA  
 TITLE JOURNAL  
 FEATURES  
 source  
 1. 1280  
 /organism="Hepatitis C virus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:11103"  
 /clone="#09-40"  
 /country="USA"  
 /note="genotype: 1b"  
 <1..>1280  
 /note="contains envelope proteins E1 and E2"  
 /codon\_start=1  
 /product="polyprotein"  
 /protein\_id="A094064.1"  
 /db\_xref="GI:53801783"  
 /translation="YEVNRSAGVHTNDNSISYETMDIMHTPGCVPTRENS  
 SRCVVALPTPLARNSIPTTTIRHVDLLVGAAFCSVMYVDLGSVLSQLPF  
 SPRRHIVDQNCISIPGHLSGHRMADMMNMSPTTALVSQLRIPQVVMVGA  
 HMGVLGAIYYSWMAKVLIVLLPAGVDGHTHTTGSAAANAAGLTLLPTFRGSSQ  
 KIQLVNTNGSMHINRTALNCDLSLQTLAALFYTHFSFNASGCPERRASRPINPFO

## ORIGIN

GNMPTITTEPHSGDORPCMYAPDQGTIPAAQVCPVYCTPSPVVGTDRFGAP  
 TYVMGANETDVLINNTRPQGNMFGCTGKFTKCGPCKIGVGNMLICPT  
 DCFRKHRENTYKCSGPMULTPRCIYDYPRLMH"

Query Match 84.4%; Score 934.2; DB 14; Length 1280;  
 Best Local Similarity 90.2%; Pred. No. 1e-194;  
 Matches 999; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 TCCTCCCGCTGCTGGGTAGGCTCATCTCCACGCTGGCGGCAAGAGCCGACATCCCC 60  
 DB 130 TCCTCCCGCTGCTGGGTAGGCTCATCTCCACGCTGGCGGCAAGAGCCGACATCCCC 189  
 QY 61 ACTGCGCAATATAGAGCCGCTGATTTGCTGTTGGGGGAGCTGCTGCTGCGCT 120  
 DB 190 ACTGCGCAATATAGAGCTGATTTGCTGTTGGGGGAGCTGCTGCTGCGCT 249  
 QY 121 ATGTAGGTGGGGATCTCTGCGGATCTGTTTCTGCTCTGAGCTGTTCACTTCTCG 180  
 DB 250 ATGTAGGTGGGGATCTCTGCGGATCTGTTTCTGCTCTGAGCTGTTCACTTCTCG 309  
 QY 181 CTTGCGGACATCAAGCGTACAGACTGCAATTGTTCAATCTATCCCGGCACTATCA 240  
 DB 310 CTTGCGGACATCAAGCGTACAGACTGCAATTGTTCAATCTATCCCGGCACTATCA 369  
 QY 241 GGTACCGGATGAGCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 370 GGTACCGGATGAGCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 429  
 QY 301 TCGCAGTACTCCGATATCCCAAGCTGTGTGGAATGATGATGATGATGATGATGATGATGAT 360  
 DB 430 TCGCAGTACTCCGATATCCCAAGCTGTGTGGAATGATGATGATGATGATGATGATGATGAT 489  
 QY 361 GTCTGTGGCGGCTTGTGCTATATTCATATGTGTGGGAACTGGGCTTAAGTCTTGTGTG 420  
 DB 490 GTCTGTGGCGGCTTGTGCTATATTCATATGTGTGGGAACTGGGCTTAAGTCTTGTGTG 549  
 QY 421 ATGCTACTCTTTGCGGCGGTGAGCGGGGAACCTTACACGACAGGGGGGACACACGCGCG 480  
 DB 550 ATGCTACTCTTTGCGGCGGTGAGCGGGGAACCTTACACGACAGGGGGGACACACGCGCG 609  
 QY 481 GCGGCCACAGGGCTTACATCCCTCTTCAACACTGGGCGCGGCTGAGAAATCGACTTGT 540  
 DB 610 AACGCTACAGGGCTTACATCCCTCTTCAACACTGGGCGCGGCTGAGAAATCGACTTGT 669  
 QY 541 AACACCAACGCGAGCTGGACATCAACAGACTGCTTGAATGCAATGATCTCCCTCA 600  
 DB 670 AACACCAACGCGAGCTGGACATCAACAGACTGCTTGAATGCAATGATCTCCCTCA 729  
 QY 601 ACTGGGTCCTTTCGCGCGCTGTTCTACAGCAAGGTTCAATGCGGATGCTCAAG 660  
 DB 730 ACTGGGTCCTTTCGCGCGCTGTTCTACAGCAAGGTTCAATGCGGATGCTCAAG 789  
 QY 661 CGCATGGCAGCTGCGGCCCATTTGACAGTGTGATCAGGGGTGGGATCCCATCACTTAT 720  
 DB 790 CGCATGGCAGCTGCGGCCCATTTGACAGTGTGATCAGGGGTGGGATCCCATCACTTAT 849  
 QY 721 AATGAGTCCACGGCTTGGACAGAGGCCCTATTTGCTGACATACGACACTCAACGCTGT 780  
 DB 850 ACTGAGGCCCAACGGCTTGGACAGAGGCCCTATTTGCTGACATACGACACTCAACGCTGT 909  
 QY 781 GGTATGCTGCGCGCGCTTGGACAGTGTGTGGCCAGGTATGATCTGTTCACTCAAGCCCTGT 840  
 DB 910 GGTATGCTGCGCGCGCTTGGACAGTGTGTGGCCAGGTATGATCTGTTCACTCAAGCCCTGT 969  
 QY 841 GTGTGTGGGACGACCGATCGTTTGGCGGCCCTTACGATAGGGGTGAGATGAGACG 900  
 DB 970 GTGTGTGGGACGACCGATCGTTTGGCGGCCCTTACGATAGGGGTGAGATGAGACG 1029  
 QY 901 GACGTGCTGCTTCAACACAGCGGCGGCCACGCGGCACTGTGTTCCGCTGTACATG 960  
 DB 1030 GACGTGCTGCTTCAACACAGCGGCGGCCACGCGGCACTGTGTTCCGCTGTACATG 1089

Qy	961	TTGAATAGCACC	GGGGTTTACCAAGACG	GTGGGGGGCCCCCGCGTCA	CATCGTGGGGGGGCT	1020
Dd	1090	ATGAATGGCACC	GGGGTTTACCAAGACG	GTGGGGGGCCCCCGCATG	AAATGGGGGGGGCT	1149
Qy	1021	GGCAACAACAC	TTTGATCTGCGCCCA	CGAGACTGCTTCCGGA	AGCATCCCGAGCCACTTAC	1080
Dd	1150	GGCAATACAC	CCCTGATCTGCGCCCA	CGAGACTGCTTCCGGA	AGCATCCCGAGCCACTTAC	1209
Qy	1081	ACCAATGCGG	TTGCGGGCCCTTGGTTG	1107		
Dd	1210	ACCAATGCGG	TTGCGGGCCCTTGGTTG	1236		
RESULT 12						
LOCUS	AY746701	1280 bp	mRNA	linear	VR1.10-OCT-2004	
DEFINITION	Hepatitis C virus clone #09-42 polyprotein mRNA, partial cds.					
ACCESSION	AY746701					
VERSION	AY746701.1 GI:53801784					
KEYWORDS						
SOURCE	Hepatitis C virus					
ORGANISM	Hepatitis C virus					
REFERENCE	Hepatitis C virus					
AUTHORS	Chambers T.J., Fan X., Droll D.A., Hembrador E., Slater T., Nickells M.W., Dustin L.B. and DiBiasegiele A.M.					
TITLE	E1/E2 quasispecies heterogeneity as a pretreatment variable during pegylated interferon therapy of chronic HCV infection					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 1280)					
AUTHORS	Chambers T.J., Fan X., Droll D.A., Hembrador E., Slater T., Nickells M.W., Dustin L.B. and DiBiasegiele A.M.					
TITLE	Direct Submission					
JOURNAL	Submitted (10-SEP-2004) Molecular Microbiology and Immunology, Saint Louis University School of Medicine, 1402 South Grand Avenue, St. Louis, MO 63104, USA					
FEATURES						
source	1..1280					
	Location/Qualifiers					
	/organism="Hepatitis C virus"					
	/mol_type="mRNA"					
	/db_xref="taxon:11103"					
	/clone="#09-42"					
	/country="USA"					
	/note="genotype: 1b"					
	<1..>1280					
	/note="contains envelope proteins E1 and E2"					
	/codon_start=1					
	/product="polyprotein"					
	/protein_id="AAU94065.1"					
	/db_xref="GI:53801785"					
	/translation="TEVRNNSAGHAYVTMDCSNSIYVETMDMINTRECVCPVREAGNS SRCAVLPPTLAARNASIPTTTRRHVLLVGAAPFCSVMVVDLCSGVFLVSQLPTFS SPRRHETVADQNCISYIPKTLGHRMAMDMNNMNSPTTALVVSQILRIPQAVDVAAG HMGVAGLAAYYSVMGNMAKVLIVMLFAGVDGHTTTGSAARAAYGLTFTGSPG KIQLVNTGSMHINTFALNCNDSLQGTFLAALFYTHSPNAGCPEPRAPIEPAQ GMGPIYTPHPGSDORPCYCMHYPADLPQCTTIAADVCGVGCFPSPVVYCTDTRGAB TYNKAETDVLILNNTTRPPQGNMFGCTMMANGTGTTKCGPCKIGVGNNTLICPT DCFRHPKATYTYKCGSPMLTPRCLIVDPYRLMH"					
ORIGIN						
Query Match	84.4% Score 934.2; DB 14; Length 1280;					
Best Local Similarity	90.2%; Pred. No. 1e-194;					
Matches	999;	Conservative	0;	Mismatches	108;	Indels 0; Gaps 0;
Qy	1	TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCTCGCGGCCAAGACGCAAGCATCCCC	60			
Dd	130	TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCTAAGGCCAAGACGCAAGCATCCCC	189			
Qy	61	ACTGGACAATATCGACGCCACGCTGATTTGCTGCTTTGGGGCGGCTGCTTCTGCTCGCT	120			
Dd	190	ACTAGACACATATCGACGCTGACGATTTGCTGCTTTGGGGCGGCTGACATTCGCTCGCTT	249			

[illegible]

Db	376	CGTGGCCGGAGTAGAGACGGATACAGGAAGTGGCAATTGCTTCACATCTATCCCGGCGCAAGTATCA	435
Qy	241	GGTACACGGCAGTGGCTTGGGATATGATATGATATGAACCTGGTCACTGACCTTAAGACAGCCCTTAGTGGTA	300
Db	436	GGCCATGGCAATGGCTTGGGATATGATATGATATGAACCTGGTCACTGACAGCCCTTAGTGGTGG	495
Qy	301	TTCGACGATCTACCGCGATATCCACCAAGCTGTGTGTGACATGTGTGGCGGGGGCCCACTGGGGGA	360
Db	496	TTCGACGATTTGCTCCGGATATCCCAAGCCGTGTGTGATATGTGTGGGGGGGCCCACTGGGGGA	555
Qy	361	GTCTGGCGGGCCCTTGGCTTACTATTCATGTGTGGGAACTGGGGCTAAAGTCTTGGTGTG	420
Db	556	GTCTTACCGGGCCCTGGCCCTACTATTTTCATGGTGGGAACTGGGGCTAAAGTCTTGAATGTG	615
Qy	421	ATGTCTACTCTTTTGGCGGCGCTTGACGGGGAACTTACACGACAGGGGGGGACACACGGCCGC	480
Db	616	ATGTCTACTCTTTTGGCGGCGCTTGACGGGGAACTTACACATCAAGGGGGGACCGAGCGCG	675
Qy	481	GCCGCCACGGGGCTTTCATCCCTCTTTCACACTGGGGCGGGCTGAGAAATTCAGCTTGTGA	540
Db	676	AAGCCCTATAGGAGCTGTGACCTTTTTCAGTTGGGCCGTGCCAGAAATTCAGCTGTGA	735
Qy	541	AACACCAACGGCAGACTGGCAGATCAACAGACATGCTTGAATCTGAAATGACTTCCTCCAA	600
Db	736	AATACCAACGGCAGACTGGCAGATCAACAGAGCTGCCGTGAATCTGAAATGACTTCCTCCAA	795
Qy	601	ACTGGGTTCTTTCGCGCGCTGTTTTCACACGACAGGTTTCAATGCGTCCGATGCTCAGAG	660
Db	796	ACCGGGTTCTTTCGCGCGCTGTTTTCACGACACAGGTTTCAATGCGTCCGATGCGCAGAG	855
Qy	661	CGCATGGCCAGCTGCGCGGCCCATTTGACAGATGATCAGGGGGTGGGGGCCATCACTTAT	720
Db	856	CGCATGGCCAGCTGCGCGGCCCATTTGACAGATGATGATCAGGGGGTGGGGGCCATCACTTAC	915
Qy	721	AATAGTCCCAACGGCTTGGACACAGAGGCCCTTATGTCTGACATACGCACTCAACCTCAACCGTGT	780
Db	916	GCTGAGCCGGGAGACTGGACACAGAGGCCCTTATGTGTGACATACGCACTCAACCTCAACCGTGC	975
Qy	781	GGTATCGTGCCCGCGCTTGGACAGGTGTGTGGCCCAAGTATCTGTTTCACTCCAGCCCTGTT	840
Db	976	GGTATTTATACCCCGCGGCGCAGGTGTGTGGCGTTCAGTGTATGCTTTCACCCCAAGCCCTGTT	1035
Qy	841	GTGTTGGGGAGCAGACCAATCGTTTCGGGGCCCTTACGTATCAAGATGGGGTGAGATGAGACG	900
Db	1036	GTGTTGGGGAGCAGACCAATCGTTTCGGGGCCCTTACGTATCAACTGGGGGGACATATGAGACG	1095
Qy	901	GACGTGCTGCTTTCACACCAACACGCGGCGGCCACAGGGGACAACTGGTTCCGCTGTATACATGG	960
Db	1096	GACGTGCTGCTTTCACACACACGCGGCGGCCCGCGGGGACAACTGGTTCCGCTGTATACATGG	1155
Qy	961	ATGAAATAGCACCAGGTTTCAACCAAGACGTGTGTGGGGGCCCCCGGTGCAACATCGGGGGGGTTC	1020
Db	1156	ATGAAATAGCACCAGGTTTCAACCAAGACGTGTGGGGGCCCCCGGTGCAACATCGGGGGGGTTC	1215
Qy	1021	GGCAACAACACTTGTGATCTGCGCCCAAGGACTGTCTCCGGAAGCATCCCGAGGCCACTTAC	1088
Db	1216	GGTAAATACACCCCTGACCTGCCCCACGAGACTGTCTCCGGAACACCCCGAGGCCACTTAC	1275
Qy	1081	ACCAAAATGCGGTTTCGGGGCCCTTGGTTG	1107
Db	1276	ACCAAAATGCGGTTTCGGGGCCCTTGGCTG	1302

RESULT 14

AY545953

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AY545953

1734 bp

DNA

linear

SYN 28-JUN-2004

Synthetic construct OH8.1 E182 region of HCV polyprotein gene,

complete cds.

AY545953

AY545953.1

GI:46403724

synthetic construct

synthetic construct



JOURNAL J. Virol. 78 (16), 8496-8505 (2004)  
 PUBMED 15280458  
 REFERENCE 2 (bases 1 to 1734)  
 AUTHORS McKeating, J.A., Zhang, L., Logvinoff, C., Flint, M., Zhang, J., Yu, J.,  
 Butera, D., Ho, D.D., Distin, L.B., Rice, C.M. and Balfe, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-FEB-2004) The Rockefeller University, Center for the  
 Study of Hepatitis C, 1230 York Ave, New York, NY 10021, USA  
 COMMENT The first 23 and last 36 bases of this set of sequences are from  
 the synthetic oligonucleotides used to generate these PCR products.  
 Neither the start nor stop codons are naturally present in HCV.  
 LOCATION/Qualifiers  
 1. 1734  
 /organism="synthetic construct"  
 /mol\_type="other DNA"  
 /db\_xref="taxon:32630"  
 /clone="OH8.3"  
 /focus  
 24. 1697  
 /organism="Hepatitis C virus"  
 /mol\_type="other DNA"  
 /db\_xref="taxon:11103"  
 1. 1734  
 /codon\_start=1  
 /transl\_table=1  
 /product="E1E2 region of HCV polyprotein"  
 /protein\_id="AA592892.1"  
 /db\_xref="GI:46403721"  
 /translation="MGCSFSLFLALLSCLTIPASAEVHNVSQVYHVTNDCSNASIV  
 YEADLIHMTPGPCVRENRSSRCWALPPTLAARNASVPTTIRHVDLGVAF  
 CSMAVYDLCGSLVLSQLPSPRHEVTDNCSIVPHVSGTHHAMDMMNMSPT  
 TALVYDQLRIPOAVDMVAGAHGVLGAAVYSMVGNKXVLIWILPFGVGEYV  
 SGTARNAVGLSTPFSVBSQIKQLVNTNGSHNRTALNCDSIQOTGLALPFAH  
 RENSISGCPERMSGRPIDKPDQGMPTTAEPSSDORPCMHAFRPGCIVAAQV  
 GPVYCTPSXVVGITDRFQVPTYNWMDNETDVLNTRPFRGNFGCTNMNSTET  
 KTCGSPNCNIGVGNVTLCTPDCKFRKHATYTKGSGWLTPRCWDVDPYLMHYP  
 CTYVFTIFKRVNMGVGEHRLTAACNMTREARNLEDRDSELSPLISTEVOVPC  
 STTLPALSTLIHONIVDQVLYGIGSAVVSFAIKMEYVLLFLALDAVACL  
 NMMLIAQAEAL"  
 1. 23  
 complement(1698..1734)  
 ORIGIN  
 primer\_bind  
 primer\_bind  
 Query Match 84.3%; Score 933.2; DB 12; Length 1734;  
 Best Local Similarity 90.2%; Pred. No. 1.6e-194;  
 Matches 998; Conservative 0; Mismatches 109; Indels 0; Gaps 0;  
 QY 1 TCCTCCGCTGCTGGGTAGGCTCACTCCGACGCTCGGCGCAAGACGCGACATCCCG 60  
 DB 196 TCCTCCGCTGCTGGGTAGGCTCACTCCGACGCTCGGCGCAAGACGCTAGCGTCCC 255  
 QY 61 ACTGCACAATAGCAGCGCACGTCGATTGCTGTTGGGGGGGCTGCTTGTCTCGCT 120  
 DB 256 ACCACGACAATAGCAGCGCACGTCGATTGCTGTTGGGGGGGCTGCTTGTCTCGCT 315  
 QY 121 ATGTAGTGGGGGATCTTCCGGGATCTGTTTCTGCTCTCAAGCTTTCACTTCTCG 180  
 DB 316 ATGTAGTGGGAGATCTTCCGGGATCTGTTTCTGCTCTCAAGCTTTCACTTCTCG 375  
 QY 181 CCTCGCGCATGAGACGCTGACGACTGCAATTTGTTCAATCTATCCCGGCGACGATCA 240  
 DB 376 CCTCGCGCGCATGAGACGCTGACGACTGCAATTTGCTCCATCTATCCCGGCGACGATCA 435  
 QY 241 GGTACACCGCATGGCTTGGGATGATGATGATGAACCTGTGACCTTAACAGACGCTTATG 300  
 DB 436 GGTACATCGCATGGCTTGGGATGATGATGATGAACCTGTGACCTTAACAGACGCTTATG 495  
 QY 301 TCGCACTACTCGGATCCCAAGCTTCTGTGAGACATGTTGGCGGGGGCCACTGGGGA 360  
 DB 496 TCGCACTTGTCTCGGATCCCAAGCGCTTGTGATGATGTTGGCGGGGGCCACTGGGGA 555  
 QY 361 GTCTTGCGGGGCTTCTACTTATTCATGTTGGGGAACCTGGGCTAAGCTTGTGTTGTG 420

DB 556 GTCTTACGGGCGCTTGCTTACTATTCATGTTGGGGAACCTGGGCTAAGCTTGTATTTGTG 615  
 QY 421 ATGCTACTCTTTCGCGCGCTTGAAGGGGAACCTTAACAGACAGGGGGGACACAGCGCGC 480  
 DB 616 ATGCTACTCTTTCGCGCGCTTGAAGGGGAACCTTAACAGACAGGGGGGACAGCGCGC 675  
 QY 481 GCGGCCACGGGCTTACATCCCTTTCACACCTGCGGCGGCTGAGAAATACGACTTGTGA 540  
 DB 676 AAGCCTTAGGGCTGTGACACTTCTTTTCACTGTTGGGCGCGCCAGAAAATACGACTGTGA 735  
 QY 541 AACACCAACGGGCTGAGCATCAACAGACTGCTTGAATGACATGACTCTCTTCAA 600  
 DB 736 AATACCAACGGGCTGAGCATCAACAGACTGCTTGAATGACATGACTCTCTTCAA 795  
 QY 601 ACTGGTCTTTCGCGCGCTTCTTACAGCACAAGTTCAATGCGTCCGATCTCAGAG 660  
 DB 796 ACCGGGTTCTTTCGCGCGCTTCTTACAGCACAAGTTCAATGCGTCCGATCTCAGAG 855  
 QY 661 CGCATGCGGAGCTGCGCGCCCATTTGACAGTTGATCAGGGGTGGGGTCCCATCACTTAT 720  
 DB 856 CGCATGCGGAGCTGCGCGCCCATTTGACAGTTGATCAGGGGTGGGGTCCCATCACTTAC 915  
 QY 721 AATGAGTCCCAAGGCTTGAACAGAGGCTTATGCTGGACATGACCACTCAACGCTGT 780  
 DB 916 GCTGAGCGGGCAGCTCGGACAGAGGCTTACTGTTGGCACTAGCACCCTGACCGTGC 975  
 QY 781 GGTATGTGCGCGCGCTTGAAGGTGTGGCCCAAGTGTACTGTTTCACTCCAAACCTGTT 840  
 DB 976 GGTATGTGACCGCGCGGCGGAGGTGTGGCTGCAAGTATGTTGCTTCAACCCCAACGCTGTT 1035  
 QY 841 GTGGGGGACGACCGCATGCTTTCGCGCGCCCTTACGTAACGATGAGGAGTGAATGAGACG 900  
 DB 1036 GTGGGGGACGACCGCATGCTTTCGCGCGCCCTTACGTAACGATGAGGAGTGAATGAGACG 1095  
 QY 901 GACGTCTGCTTCTCAACAACAGCGCGCGCCCAAGGAGCAACTGTTGCGCTGTACATGG 960  
 DB 1096 GACGTCTGCTTCTTAACAACAAGCGCGCGCGCGGAGCAACTGTTGCGCTGTACATGG 1155  
 QY 961 ATGAATAGCACCGGGTTTCAACAAGACGTGTGGGGGGCCCCCGTGTCAACATCGGGGGGCTC 1020  
 DB 1156 ATGAATAGCACCTGGGTTTCAACAAGACGTGTGGGGGGCCCCCGTGTCAACATCGGGGGGCTC 1215  
 QY 1021 GGCACAACACTTGTATGCTGCGCCAGAGACTGCTTCGGAAGCATCCGAGGCACTTAC 1080  
 DB 1216 GGTATATACACTCTGTGACTGCGCCACGAGACTGCTTCGGAACACCCCGAGGCACTTAC 1275  
 QY 1081 ACCAATGCGGTTTGGGGGCTTGGTTG 1107  
 DB 1276 ACCAATGCGGTTTGGGGGCTTGGCTG 1302

Search completed: February 21, 2005, 07:50:41  
 Job time : 5150 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 04:05:51 ; Search time 702 Seconds  
(without alignments)  
9334.977 Million cell updates/sec

Title: US-09-664-363-19

Sequence: 1 TCTCTCCGCTGCTGGTAGC.....GCGGTCGGGCGCTTGGTTG 1107

Scoring table: IDENTITY\_NUC

Searched: 4390206 beqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

1: geneseqn\_1edeca4: \*  
2: geneseqn1980s: \*  
3: geneseqn1990s: \*  
4: geneseqn2000as: \*  
5: geneseqn2001as: \*  
6: geneseqn2001bs: \*  
7: geneseqn2002as: \*  
8: geneseqn2002bs: \*  
9: geneseqn2003as: \*  
10: geneseqn2003bs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004as: \*  
13: geneseqn2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1107	100.0	1107	2	AAQ12240	AaQ12240 Clone 135
2	1107	100.0	2116	2	AAQ12242	AaQ12242 Encodes F
3	927.8	83.8	9605	6	ABK91431	AbK91431 Hepatitis
4	927.8	83.8	9605	6	ABK91424	AbK91424 Hepatitis
5	927.8	83.8	9605	6	ABK91429	AbK91429 Hepatitis
6	927.8	83.8	9605	6	ABK91432	AbK91432 Hepatitis
7	927.8	83.8	9605	6	ABK91411	AbK91411 Hepatitis
8	927.8	83.8	9605	6	ABK91430	AbK91430 Hepatitis
9	927.8	83.8	9605	6	ABK91428	AbK91428 Hepatitis
10	927.8	83.8	9605	6	ABK91425	AbK91425 Hepatitis
11	927.8	83.8	9605	6	ABK91426	AbK91426 Hepatitis
12	927.8	83.8	9605	6	ABK91433	AbK91433 Hepatitis
13	927.8	83.8	9605	6	AAD25332	Aad25332 Hepatitis
14	927.8	83.8	9608	6	ABK91427	AbK91427 Hepatitis
15	927.8	83.8	11062	3	AAD25331	Aad25331 Hepatitis
16	927.8	83.8	11076	3	AAA98965	AAa98965 Hepatitis
17	919.4	83.1	2187	2	ABA03491	AbA03491 Citiclic F
18	919.4	83.1	2540	2	AAQ34889	AaQ34889 NMBV
19	919.4	83.1	2540	2	AAQ65753	AaQ65753 NMBV 9C
20	916.2	82.8	2540	2	AAQ26628	AaQ26628 Hepatitis

21	912.5	82.4	9436	2	AA063399	Ad63399 Blood tra
22	909.8	82.2	9555	2	AA024643	Ad24643 Infection
23	909.8	82.2	9555	4	AA023492	Ad23492 Infection
24	909.8	82.2	9555	4	AA086639	Ad86639 Nucleotid
25	909.8	82.2	9555	12	AA083222	Ad083222 Hepatitis
26	909.8	82.2	9555	12	AA079336	Ad079336 Hepatitis
27	909.8	82.2	9559	2	AA024633	Ad24633 Infection
28	908.2	82.0	2187	2	AB003492	Ab003492 Cuticle p
29	904.2	81.7	1953	8	AA055222	Ad155222 Plasmid p
30	904.2	81.7	2829	2	AA060673	Ad60673 Fragment
31	903.8	81.6	3461	2	AA060468	Ad60468 Non-A, no
32	903.8	81.6	3461	2	AA030386	Ad30386 5'UTR/COR
33	902.2	81.5	9587	13	ADR82189	Ad82189 Hepatitis
34	899	81.2	9402	2	AA041345	Ad41345 Human hep
35	897.4	81.1	9405	2	AA040426	Ad40426 Full-leng
36	895.8	80.9	9400	13	ADR82190	Ad82190 Hepatitis
37	893.2	80.7	9414	13	ADR82190	Ad82190 Hepatitis
38	893.2	80.7	9472	2	AA033882	Ad33882 Korean he
39	892.6	80.6	2238	12	ADJ82982	Adj82982 Adenovect
40	892.6	80.6	3081	12	ADJ82983	Adj82983 Adenovect
41	892.6	80.6	37035	12	ADJ82985	Adj82985 Adenovect
42	891	80.5	33650	2	AA0703677	Ad0703677 Hepatitis
43	891	80.5	7911	2	AA032436	Adg32436 HCV anti
44	891	80.5	8460	12	ADN11771	Adn11771 Hepatitis
45	891	80.5	9413	2	AA080498	Ad80498 DNA encod

## ALIGNMENTS

RESULT 1  
AAQ12240  
ID AAQ12240 standard; DNA; 1107 BP.

DT	25-MAR-2003	(revised)
DT	06-SEP-1991	(first entry)

DE Clone 136/155 encoding PT-NANBH viral structural protein.

KW post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.

OS Non-A.

OS non-B hepatitis virus.

PN GB2239245-A.

PD 26-JUN-1991.

PF 17-DEC-1990; 90GB-00027250.

PR 18-DEC-1989; 89GB-00028562.

PR 03-MAR-1990; 90GB-00004814.

PA (WELL ) WELLCOME FOUND LTD.

2 XX

XX

DR P-PSDB; AAR12598.

PT Post-transfusion r

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841

XXXXXXXXXXXXXXXXXXXX

CC viral genome which are antigenic. It was isolated from serum of humans

CC on 25-MAR-2003 to correct PA field.)

XX Sequence 1107 BP; 205 A; 344 C; 319 G; 239 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 1107; DB 2; Length 1107;  
Best Local Similarity 100.0%; Pred. No. 5,6e-308;  
Matches 1107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCGCCGCAAGAGCGCAGCATCCCC 60  
DB 1 TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCGCCGCAAGAGCGCAGCATCCCC 60  
QY 61 ACTGCCAATAATGAGAGCGCCAGTCGATTTGCTGTTGGGGCGGCTGCTTCTGCTCGCT 120  
DB 61 ACTGCCAATAATGAGAGCGCCAGTCGATTTGCTGTTGGGGCGGCTGCTTCTGCTCGCT 120  
QY 121 ATGTAGTGGGGGATCTGTGGGATCGTGTCTCTGCTCTAGCGTTCACTTCCTG 180  
DB 121 ATGTAGTGGGGGATCTGTGGGATCTGTGTCTCTGCTCTAGCGTTCACTTCCTG 180  
QY 181 CCTGCGCAGATCAGAGCGGTACAGACTGCAATTTGTCATCTATCCCGCCAGATACA 240  
DB 181 CCTGCGCAGATCAGAGCGGTACAGACTGCAATTTGTCATCTATCCCGCCAGATACA 240  
QY 241 GGTCAACCGCATGCGCTTGGGATATGATGATGAATGCTGTCACTTACAGACGCTTATG 300  
DB 241 GGTCAACCGCATGCGCTTGGGATATGATGATGAATGCTGTCACTTACAGACGCTTATG 300  
QY 301 TCGCAGCTACTCGGGATCCCAAGAGCTGTGGAACAATGGGGGGGCGCACTGGGGA 360  
DB 301 TCGCAGCTACTCGGGATCCCAAGAGCTGTGGAACAATGGGGGGGCGCACTGGGGA 360  
QY 361 GTCCTGGCGGCGCTTGGCTTACTATTCATAGTGGGGAATGAGTCTTGTGTG 420  
DB 361 GTCCTGGCGGCGCTTGGCTTACTATTCATAGTGGGGAATGAGTCTTGTGTG 420  
QY 421 ATGCTACTCTTTGCGGCGCTTGAACGGGAACTTTACAGACAGGGGGGACACACGGCGC 480  
DB 421 ATGCTACTCTTTGCGGCGCTTGAACGGGAACTTTACAGACAGGGGGGACACACGGCGC 480  
QY 481 GCGGCCACGGGGCTTTCATCCCTTTCACACCTGGGGCGGCTCAGAAAATCCAGCTTGT 540  
DB 481 GCGGCCACGGGGCTTTCATCCCTTTCACACCTGGGGCGGCTCAGAAAATCCAGCTTGT 540  
QY 541 AACACCAACGGCAGCTGGCAGATCAACAGAACTGCTTGAATGACATGATCCTCTCA 600  
DB 541 AACACCAACGGCAGCTGGCAGATCAACAGAACTGCTTGAATGACATGATCCTCTCA 600  
QY 601 ACTGGGTTCTTGGCGGCTGTTCTACAGCAAGTTCAATGGCTCCGATGCTCAGAG 660  
DB 601 ACTGGGTTCTTGGCGGCTGTTCTACAGCAAGTTCAATGGCTCCGATGCTCAGAG 660  
QY 661 CGCATGGCCAGCTGGCGGCCCATTTGACAGTGGATAGGGGGTGGGATCCCATCACTAT 720  
DB 661 CGCATGGCCAGCTGGCGGCCCATTTGACAGTGGATAGGGGGTGGGATCCCATCACTAT 720  
QY 721 AATGAGTCCACAGCGCTTGAACAGAGGCGCTTATTTGCTGACATACGACCTCAACCTGT 780  
DB 721 AATGAGTCCACAGCGCTTGAACAGAGGCGCTTATTTGCTGACATACGACCTCAACCTGT 780  
QY 781 GGTATCGTCCCGCGTGTGACAGTGTGTGCCCCAGTGTACTGTTCACTCCAAAGCCCTGT 840  
DB 781 GGTATCGTCCCGCGTGTGACAGTGTGTGCCCCAGTGTACTGTTCACTCCAAAGCCCTGT 840  
QY 841 GTGTGGGGAGACACCGATCGTTTGGGGCGCCCTTACGTAAGATGGGTGGAATGAGACG 900  
DB 841 GTGTGGGGAGACACCGATCGTTTGGGGCGCCCTTACGTAAGATGGGTGGAATGAGACG 900  
QY 901 GACGTGTGCTTCTCAACAACACCGCGCCGCAAGGGGCAATGTGCTGCTGTACATGG 960  
DB 901 GACGTGTGCTTCTCAACAACACCGCGCCGCAAGGGGCAATGTGCTGCTGTACATGG 960  
QY 961 ATGAATGACACCGGCTTCAACCAAGACGTGTGGGGGCGCCCGTGTCAACATCGGGGGGTC 1020  
DB 961 ATGAATGACACCGGCTTCAACCAAGACGTGTGGGGGCGCCCGTGTCAACATCGGGGGGTC 1020

DB 961 ATGAATGACACCGGCTTCAACCAAGACGTGTGGGGGCGCCCGTGTCAACATCGGGGGGTC 1020  
QY 1021 GGCAACAACACTTTGATCTGCCCAAGACTGCTTCCGAGACATCCGAGGCCACTTAC 1080  
DB 1021 GGCAACAACACTTTGATCTGCCCAAGACTGCTTCCGAGACATCCGAGGCCACTTAC 1080  
QY 1081 ACCAAATGCGGTTCCGGGCGCTTGGTTG 1107  
DB 1081 ACCAAATGCGGTTCCGGGCGCTTGGTTG 1107  
RESULT 2  
AAQ12242  
ID AAQ12242 standard; DNA; 2116 BP.  
XX  
AC AAQ12242;  
XX  
DT 25-MAR-2003 (revised)  
DT 17-SEP-1991 (first entry)  
XX  
DE Encodes PT-NANBH viral structural and non-structural proteins.  
XX  
KM post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.  
OS Non-A.  
OS non-B hepatitis virus.  
XX  
FH Key Location/Qualifiers  
FT CDS 108..2116  
FT /tag= a  
XX  
PN GB2239245-A.  
XX  
PD 26-JUN-1991.  
XX  
PF 17-DEC-1990; 90GB-00027250.  
XX  
PR 18-DEC-1989; 89GB-00028562.  
PR 27-FEB-1990; 90GB-00004414.  
PR 03-MAR-1990; 90GB-00004814.  
XX  
PA (WELL) WELLCOME FOUND LTD.  
PA (HIGH/) HIGHFIELD P E.  
XX  
PI Highfield PE, Rodgers BC, Tedder RS, Barbara JW;  
XX  
DR WPI; 1991-187584/26.  
XX  
DR P-PSDB; AAR12600.  
XX  
PT Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA  
XX and antibodies used in diagnostic assays and in vaccines.  
PS Claim 10; Page 83-87; 108pp; English.  
XX  
XX This sequence is thought to encode viral structural and non- structural  
CC proteins of the PT-NANBH viral genome which are antigenic. It was  
CC isolated from human serum infectious for the virus. See also AAQ12236-41.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 2116 BP; 392 A; 650 C; 624 G; 450 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1107; DB 2; Length 2116;  
Best Local Similarity 100.0%; Pred. No. 7.2e-308;  
Matches 1107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCTTGGCGCCCAAGAGCGCAGCATCCCC 60  
DB 1010 TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCTTGGCGCCCAAGAGCGCAGCATCCCC 1069  
QY 61 ACTGGCAATAATGAGAGCGCCAGTCGATTTGCTGTTGGGGCGGCTGCTTCTGCTCGCT 120  
DB 1070 ACTGGCAATAATGAGAGCGCCAGTCGATTTGCTGTTGGGGCGGCTGCTTCTGCTCGCT 1129

```

QY 121 ATGTACGTGGGGATCTCTGCGGATCTGTTTCTCTGCTCTCAGCTGTTCACCTTCTCG 180
  |||
DB 1130 ATGTACGTGGGGATCTCTGCGGATCTGTTTCTCTGCTCTCAGCTGTTCACCTTCTCG 1189
QY 181 CCTCGCGACATCAGACGGGTACAGACTGTCAATTTGTTCAATCTATCCCGGCACGTATCA 240
  |||
DB 1190 CCTCGCGACATCAGACGGGTACAGACTGTCAATTTGTTCAATCTATCCCGGCACGTATCA 1249
QY 241 GGTACACCGCATGGCTTGGATATGATGATGAACTGGTCACTTACAGACGCCCTTAATGTA 300
  |||
DB 1250 GGTACACCGCATGGCTTGGATATGATGATGAACTGGTCACTTACAGACGCCCTTAATGTA 1309
QY 301 TCGACGCTACTCGGATCCCAAGGCTGTGTGACATGTGTGCGGGGGGCCCACTGAGGA 360
  |||
DB 1310 TCGACGCTACTCGGATCCCAAGGCTGTGTGACATGTGTGCGGGGGGCCCACTGAGGA 1369
QY 361 GTCTGTGGCGGCTTGTCTTCAATTCATGTTGGGAACTGGGCTTAAGTCTTGGTTGTG 420
  |||
DB 1370 GTCTGTGGCGGCTTGTCTTCAATTCATGTTGGGAACTGGGCTTAAGTCTTGGTTGTG 1429
QY 421 ATGCTACTCTTTTCCGGCGCTTGAAGGGGAACTTTACACAGAGGGGGGACACACGGCCGC 480
  |||
DB 1430 ATGCTACTCTTTTCCGGCGCTTGAAGGGGAACTTTACACAGAGGGGGGACACACGGCCGC 1489
QY 481 GCGGCCACAGGGGCTTACATCCCTCTTCAACACTGGGCGGCTCAGAAATCCAGCTTGT 540
  |||
DB 1490 GCGGCCACAGGGGCTTACATCCCTCTTCAACACTGGGCGGCTCAGAAATCCAGCTTGT 1549
QY 541 AACACCAAGGAGCTGAGCATCAACAGAACTGCTTGAATGCAATGAATCCCTCCCA 600
  |||
DB 1550 AACACCAAGGAGCTGAGCATCAACAGAACTGCTTGAATGCAATGAATCCCTCCCA 1609
QY 601 ACTGGGTTCTTCCCGCGCTGTCTTACAGCAGCAAGTTCAATGCTCCGATGCTCAGAG 660
  |||
DB 1610 ACTGGGTTCTTCCCGCGCTGTCTTACAGCAGCAAGTTCAATGCTCCGATGCTCAGAG 1669
QY 661 CGATGAGCAGCTGCGGCCCATTTGACAGTTGATCAGGGGTGGGCTCCCATCACTTAT 720
  |||
DB 1670 CGATGAGCAGCTGCGGCCCATTTGACAGTTGATCAGGGGTGGGCTCCCATCACTTAT 1729
QY 721 AATGATGCCACGGCTTGGACAGAGGCCCTTATTTGTGTGACATACGCACTCAACCGTGT 780
  |||
DB 1730 AATGATGCCACGGCTTGGACAGAGGCCCTTATTTGTGTGACATACGCACTCAACCGTGT 1789
QY 781 GGTATGTCGCGCGCTTGTGACAGTGTGTGCGCAGTGTACTGTTTCAAGCCCTGTT 840
  |||
DB 1790 GGTATGTCGCGCGCTTGTGACAGTGTGTGCGCAGTGTACTGTTTCAAGCCCTGTT 1849
QY 841 GTGATGGGAGCAGCCGATCGTTTCCGCGCCCTTACGATGAGTGGGTGGAATGAGAGC 900
  |||
DB 1850 GTGATGGGAGCAGCCGATCGTTTCCGCGCCCTTACGATGAGTGGGTGGAATGAGAGC 1909
QY 901 GACGTCTGCTTCTCAACAACACGGCGCCGACAGGGGGCAATGTTCCGCTGTACATG 960
  |||
DB 1910 GACGTCTGCTTCTCAACAACACGGCGCCGACAGGGGGCAATGTTCCGCTGTACATG 1969
QY 961 ATGAATAGCACCGGTTTACCAAGAGCTGTGTGGGGGCCCCCGGTGCAATCGGGGGGCT 1020
  |||
DB 1970 ATGAATAGCACCGGTTTACCAAGAGCTGTGTGGGGGCCCCCGGTGCAATCGGGGGGCT 2029
QY 1021 GGAACAACACTTGTATGTGCGCCACAGAGTGTCTTCCGGAAGATCCCGAGGCACTTAC 1080
  |||
DB 2030 GGAACAACACTTGTATGTGCGCCACAGAGTGTCTTCCGGAAGATCCCGAGGCACTTAC 2089
QY 1081 ACCAAATGCGGTTCCGGGGCTTGGTTG 1107
  |||
DB 2090 ACCAAATGCGGTTCCGGGGCTTGGTTG 2116

```

RESULT 3  
ABK91431  
ID ABK91431 standard; DNA; 9605 BP.  
XX

```

AC ABK91431;
XX
XX 15-NOV-2002 (first entry)
DE Hepatitis C virus Con 1 isolate DNA mutant 8.
XX
XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KM internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 342..9374
FT CDS
FT
FT /tag= a
FT /product= "HCV polypeptide"
FT /note= "The polypeptide consists of the Core, E1, E2, p7,
FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT replace(6934,T)
FT mutation
FT
FT /tag= b
PN W0200259321-A2.
XX
XX 01-AUG-2002.
PD
XX
XX 16-JUN-2002; 2002WC-EP000526.
XX
XX 23-JUN-2001; 2001US-0263479P.
XX
XX (RICE-) IST RICEGHE BIOL MOLECOLARE ANGELETTI.
PI De Francesco R, Migliaccio G, Paonessa G;
XX
XX WPI; 2002-599793/4.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX PT ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX
XX Claim 9; Page; 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX CC vector comprising a nucleotide sequence coding for the altered nucleic
XX CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX CC recombinant cell human hepatoma cell comprising the altered nucleic acids
XX CC; (3) a recombinant cell produced by introducing into a human hepatoma
XX CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX CC replicon enhanced cell or which containing a functional HCV replicon; (5)
XX CC an HCV replicon enhanced cells made in the method; and (6) measuring the
XX CC ability of a compound to affect HCV activity. The HCV replicons and HCV
XX CC replicon enhanced cells are useful in studying HCV replication and
XX CC expression, and HCV and host cell interactions, producing HCV RNA and
XX CC proteins, and providing a system for measuring the ability of a compound
XX CC to modulate one or more HCV activities e.g. to discover drugs which may
XX CC treat HCV mediated diseases such as liver failure, cirrhosis and
XX CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
XX CC mutant of the invention. Note: The present sequence is not shown in the
XX CC specification but was created by the indexer using the HCV sequence
XX CC appearing as ABK91431 and the information in Claim 9
XX
SQ Sequence 9605 BP; 1910 A; 2884 C; 2733 G; 2078 T; 0 U; 0 Other;

```

Query Match 83.8%; Score 927.8; DB 6; Length 9605;  
Best Local Similarity 89.9%; Pred. No. 5; le-256;  
Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
QY 1 TCCTCCGCTGCTGGGAGGCTCCTCCACGCTGCGGCGCAAGAGCGCCACATCCCC 60

Db 1044 TCCTCCCGCTGCTGGTAGGCGCTCACTCCACGCTCGCGGCAGAACTGACGCTCC 1103  
QY 61 ACTGCGCAATATAGAGCGCCAGTGCATTTGCTGTTGGGGGGGCTGCTTCTGCTCCGCT 120  
Db 1104 ACTACGACGATAGAGCGCCAGTGTGATTTTCTGTGGGGGGGCTGCTTCTGCTCCGCT 1163  
QY 121 ATGTAGCTGGGGGATCTCTGCGGATCTGTTTCTCTCTCTGAGCTGTTCACTTCTCG 180  
Db 1164 ATGTACGTGGGAGATCTCTGCGGATCTGTTTCTCTGCTGCGGAGCTTCACTTCTCG 1223  
QY 181 CCTGCGGACATAGAGCGGTACAGACTGCAATTTGTTCAATCTATCCCGGCACGATCA 240  
Db 1224 CCTGCGCGGACAGAGACAGTACAGACTGCAATTTGTTCAATCTATCCCGGCACGATCA 1283  
QY 241 GGTACCGGATGGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
Db 1284 GGTACCGGATGGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1343  
QY 301 TCGCAGCTACTCCGATCCCAAGCTGTGTGAGCAATGATGGCGGGGCCCACTGGGGA 360  
Db 1344 TCGCAGTACTCCGATCCCAAGCTGTGTGAGCAATGATGGCGGGGCCCACTGGGGA 1403  
QY 361 GTCTGTGGCGGCTTGTCTTCTATTTCCATGTGGGGAACCTGGGTTAAGTCTTGGTTGTG 420  
Db 1404 GTCTGTGGCGGCTTGTCTTCTATTTCCATGTGGGGAACCTGGGTTAAGTCTTGGTTGTG 1463  
QY 421 ATGCTCTCTTTGCGCGGCTGTGACGGGGAACTTTACAGCAAGGGGGGACACACGGGCGC 480  
Db 1464 ATGCTCTCTTTGCGCGGCTGTGACGGGGAACTTTATGTGACAGGGGGGACAGATGGCCAA 1523  
QY 481 GCCGCCACGGGCTTACATCTCTTTCACACTGTGGCGGGCTGAGAAATCCAGCTTGTGA 540  
Db 1524 AACACCTCGGGATTAGTCCCTTTTTCACCCGGGCTCATCCAGAAAATCCAGCTTGTGA 1583  
QY 541 AACACCAACGGGCTGTGGCAATCAACAGAACTGCTTGAATGCAATGCTCTTCCAA 600  
Db 1584 AACACCAACGGGCTGTGGCAATCAACAGAACTGCTTGAATGCAATGCTCTTCCAA 1643  
QY 601 ACTGGGTTCTTTCGCGGCTGTGTACAGCAAGGTTCAATGTGCGGATGCTCAGAG 660  
Db 1644 ACTGGGTTCTTTCGCGGCTGTGTGTACAGCAAGTTCAATGTGATGATGCTCAGAG 1703  
QY 661 CGCATGGCGAGTCCGCGCCCATTTGACAGTTCAGTTCAGGGGTGGGCTCCCATCTTAT 720  
Db 1704 CGCATGGCGAGTCCGCGCCCATTTGACAGTTCAGTTCAGGGGTGGGCTCCCATCTTAC 1763  
QY 721 AATGATGCCAGCGCTTGTGACAGAGGCTTATTTGCTGCGACTTACGCACTTCAACGCTGT 780  
Db 1764 AATGATGCCAGCGCTTGTGACAGAGGCTTATTTGCTGCGACTTACGCACTTCAACGCTGT 1823  
QY 781 GGTATCGTGCCTGGGTGTGAGGTGTGTGAGGCGGAGTACGTTTCACTTCAAGGCTTGT 840  
Db 1824 GGTATCGTGCCTGGGTGTGAGGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1883  
QY 841 GTGTGTGGGAGAGCCGATCTTTCGCGGCTTACGATGATGATGATGATGATGATGATGAT 900  
Db 1884 GTGTGTGGGAGAGCCGATCTTTCGCGGCTTACGATGATGATGATGATGATGATGATGAT 1943  
QY 901 GACGTGTCTTCTTCAACAACAGCGGCGCGCCAGCGGGCAACTGTGTGCTGTACATGG 960  
Db 1944 GACGTGTCTTCTTCAACAACAGCGGCGCGCCAGCGGGCAACTGTGTGCTGTACATGG 2003  
QY 961 AATGATGACACGGGTTCAACAAGAGTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1020  
Db 2004 AATGATGACACGGGTTCAACAAGAGTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2063  
QY 1021 GGGCAACAACCTTGTATCTGCCCCAGAGACTGCTTCCGGAAGCATCCGAGGCACTTAC 1080  
Db 2064 GGGCAACAACCTTGTATCTGCCCCAGAGACTGCTTCCGGAAGCATCCGAGGCACTTAC 2123  
QY 1081 ACCAATGCGGTTTCGGGGCTTGTGTTG 1107

Db 2124 ACCAATGTTGTTTCGGGGCTTGTGTTG 2150  
RESULT 4  
ABK91424  
ID ABK91424 standard; DNA; 9605 BP.  
XX  
AC ABK91424;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE Hepatitis C virus Con 1 isolate DNA mutant 1.  
XX  
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
OS Hepatitis C virus.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH 342..9374  
FT CDS  
FT /\*tag= a  
FT /product= "HCV polypeptide"  
FT /note= "The polypeptide consists of the Core, E1, E2, P7,  
FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
FT mutation  
FT replace(3625,G)  
FT /\*tag= b  
XX  
PN MO200259321-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 16-JAN-2002; 2002MO-EP000526.  
XX  
PR 23-JUN-2001; 2001JUS-0263479P.  
XX  
PA (RICE-) IST RIGERCHE BIOL MOLECOLARE ANGELETTI.  
XX  
PI De Francesco R, Migliaccio G, Paonessa G;  
XX  
XX WPI; 2002-599793/64.  
XX  
XX New Hepatitis C virus (HCV) replicons comprising altered HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV)  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX  
PS Claim 9; Page; 69pp; English.  
XX  
XX The invention relates to nucleic acid molecules comprising altered HCV  
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
CC internal ribosome entry site (IRES) region coding for one or more NS3,  
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
CC are detailed in the specification. Also included are (1) an expression  
CC vector comprising a nucleotide sequence coding for the altered nucleic  
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
CC; (3) a recombinant cell produced by introducing into a human hepatoma  
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC proteins, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
CC mutant of the invention. Note: The present sequence is not shown in the  
CC specification but was created by the indexer using the HCV sequence  
XX appearing as ABK91411 and the information in Claim 9



CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
 CC mutant of the invention. Note: The present sequence is not shown in the  
 CC specification but was created by the indexer using the HCV sequence  
 CC appearing as ABK91411 and the information in Claim 9

XX Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 U; 0 Other;

Query Match 83.8%; Score 927.8; DB 6; Length 9605;

Best Local Similarity 89.9%; Pred. No. 5,1e-256;

Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```

QY 1 TCCCTCCGCTGCTGGGTAGCGCTCACTCCACCGCTCGCGGCAAGACGCGACATCCCC 60
DB 1044 TCCCTCCGCTGCTGGGTAGCGCTCACTCCACCGCTCGCGGCAAGACGCTAGGTCCTCC 1103
QY 61 ACTGCGCAATATGACGCGCCAGTGTGCTGTTGGGCGGCTGCTTCTGCTCCGCT 120
DB 1104 ACTACGACGATACGACGCGCATGTCGATTTGCTGTTGGGCGGCTGCTTCTGCTCCGCT 1163
QY 121 ATGTAGTGGGGATCTGTGCGGATCTGTTTCTGCTCTGACGCTTCACTTCTCG 180
DB 1164 ATGTAGTGGGAATCTGTGCGGATCTGTTTCTGCTGCGCCAGCTGTTCACCTTCTCG 1223
QY 181 CCTGCGGACATGACGAGTACAGACTGCAATTTGTTCAATCTATCCCGGCAAGTATCA 240
DB 1224 CCTGCGGACGACGAGTACAGACTGCAATTTGTTCAATCTATCCCGGCAAGTATCA 1283
QY 241 GGTACCGCATGCTGGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 300
DB 1284 GGTACCGCATGCTGGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1343
QY 301 TCGCAGTATCTCGGATATCCCAAGCTGTGTGGAATGATGATGATGATGATGATGATG 360
DB 1344 TCGCAGTATCTCGGATATCCCAAGCTGTGTGGAATGATGATGATGATGATGATGATG 1403
QY 361 GTCCTGCGGCGCTTGTGCTTATTTCCATGTTGGGGAATGAGTCTTGTGTTGTG 420
DB 1404 GTCCTGCGGCGCTTGTGCTTATTTCCATGTTGGGGAATGAGTCTTGTGTTGTG 1463
QY 421 ATGCTACTTTTTCGCGGCTTGAACGGGGAACCTTTACAGACAGGGGGGACACACGCGCG 480
DB 1464 ATGCTACTTTTTCGCGGCTTGAACGGGGAACCTTTACAGACAGGGGGGACATGAGGCA 1523
QY 481 GCGGCGGACGCGGCTTATCCCTTTCACACTGCGCGGCTCAGAAATCCAGCTTGT 540
DB 1524 AACACCTCGGGATATGATGCTCTTTTTCACCGGCGCTATCCAGAAATCCAGCTTGT 1583
QY 541 AACACCAACGCGGCTTATCCCTTTCACACTGCGCGGCTCAGAAATCCAGCTTGT 600
DB 1584 AACACCAACGCGGCTTATCCCTTTCACACTGCGCGGCTCAGAAATCCAGCTTGT 1643
QY 601 ACTGGGTTCTTTCGCGGCTTGTTCACGCAACAGTTCAATGCTCGGATGCTCAG 660
DB 1644 ACTGGGTTCTTTCGCGGCTTGTTCACGCAACAGTTCAATGCTCGGATGCTCAG 1703
QY 661 CGATGCGGCGGCTTGTTCACGCAACAGTTCAATGCTCGGATGCTCAG 720
DB 1704 CGATGCGGCGGCTTGTTCACGCAACAGTTCAATGCTCGGATGCTCAG 1763
QY 721 AATGATGCCAGGCTTGTTCACGCAACAGTTCAATGCTCGGATGCTCAG 780
DB 1764 AATGATGCCAGGCTTGTTCACGCAACAGTTCAATGCTCGGATGCTCAG 1823
QY 781 GGTATGCTGCTGCTGCTTGTTCAGTGTGCGGCAAGTATGTTCACTTCAAGCCCTGT 840
DB 1824 GGTATGCTGCTGCTGCTTGTTCAGTGTGCGGCAAGTATGTTCACTTCAAGCCCTGT 1883
QY 841 GGTATGCGGACGACGATGCTTGTTCAGTGTGCGGCAAGTATGTTCACTTCAAGCCCTGT 900
DB 1884 GGTATGCGGACGACGATGCTTGTTCAGTGTGCGGCAAGTATGTTCACTTCAAGCCCTGT 1943

```

```

QY 901 GAGTGTCTCTTCTCAACACACGCGCGCCGCAAGGAGGCACTGTTGAGTACATG 960
DB 1944 GAGTGTCTCTTCTTAAACACACGCGCGCCGCAAGGAGGCACTGTTGAGTACATG 2003
QY 961 ATGAATATGACCGGGTTTACCAAGAGTGTGGGGGCCCCCGTGCATATCCGGGGGTC 1020
DB 2004 ATGAATATGACCGGGTTTACCAAGAGTGTGGGGGCCCCCGTGCATATCCGGGGGTC 2063
QY 1021 GGCACACACATTTGATCTGCCCCACGAGCTGCTTCCGAGAGATCCGAGGCACTTAC 1080
DB 2064 GGCATATTAACCTTGACTGCTCCCGACGAGCTGCTTCCGAGAGATCCCGAGGCACTTAC 2123
QY 1081 ACCAAATGCGGTTTGGGGGCTTGTGTTG 1107
DB 2124 ACCAAGTGTGTTTGGGGGCTTGTGTTG 2150

```

# RESULT 6

ABK91432  
 ID ABK91432 standard; DNA; 9605 BP.

XX ABK91432;

XX AC 15-NOV-2002 (first entry)

XX DT Hepatitis C virus Con 1 isolate DNA mutant 9.

XX DE HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;

XX KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;

XX KM internal ribosome entry site; IRES; NS5A; HCV replication; mutant.

XX OS Hepatitis C virus.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT CDS 342..9374 /tag= a

FT /product= "HCV polyprotein"

FT /note= "The polyprotein consists of the Core, E1, E2, P7,

FT mutation NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"

FT /tag= b

XX WO200259321-A2.

XX PD 01-AUG-2002.

XX PF 16-JAN-2002; 2002WO-EP000526.

XX PR 23-JAN-2001; 2001US-0263479P.

XX PA (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.

XX PT De Francesco R, Migliaccio G, Paonessa G;

XX DR WPI; 2002-599793/64.

XX PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV

XX PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal

XX PT ribosome entry site (IRES) region, useful in studying HCV replication and

XX PS expression.

XX PS Claim 9; Page; 69pp; English.

XX CC The invention relates to nucleic acid molecules comprising altered HCV

XX CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)

XX CC internal ribosome entry site (IRES) region coding for one or more NS3,

CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus) replicon enhanced cell or which containing a functional HCV replicon; (5) an HCV replicon enhanced cells made in the method; and (6) measuring the ability of a compound to affect HCV activity. The HCV replicons and HCV replicon enhanced cells are useful in studying HCV replication and expression, and HCV and host cell interactions, producing HCV RNA and proteins, and providing a system for measuring the ability of a compound to modulate one or more HCV activities e.g. to discover drugs which may treat HCV mediated diseases such as liver failure, cirrhosis and hepatocellular carcinoma. The present sequence is an HCV replicon Con 1 mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the HCV sequence appearing as ABK91411 and the information in Claim 9

XX Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 U; 0 Other;

Query Match 83.8%; Score 927.8; DB 6; Length 9605;  
Best Local Similarity 89.9%; Pred. No. 5,1e-256;  
Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```

QY 1 TCCTCCCGCTGCTGGGAGGCTCACTCCACGCTCGGCGCAAGAGCGGAGATCCCC 60
DB 1044 TCCTCCCGCTGCTGGGAGGCTCACTCCACGCTCGGCGCAAGAGCGGAGATCCCC 1103
QY 61 ACTGCGCAATATGAGCGCGCAAGCTGATTTGCTGTTGGGCGGCTGCTGCTCGCT 120
DB 1104 ACTGCGCAATATGAGCGCGCAAGCTGATTTGCTGTTGGGCGGCTGCTGCTCGCT 1163
QY 121 ATGTAGCTGGGGGATCTCTGCGGATCTGTTTCTGCTCTCACTGTTTCACTTCTCG 180
DB 1164 ATGTAGCTGGGGGATCTCTGCGGATCTGTTTCTGCTCTCACTGTTTCACTTCTCG 1223
QY 181 CTTCCGCGCATAGAGCGGATGAGAGTGCATTTGTCATTTCCCGGCGAGATCA 240
DB 1224 CTTCCGCGCGAGAGAGATGAGAGTGCATTTGTCATTTCCCGGCGAGATCA 1283
QY 241 GGTACACGCGATGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 1284 GGTACACGCGATGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1343
QY 301 TCGCAGCTACTCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCC 360
DB 1344 TCGCAGCTACTCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCC 1403
QY 361 GTCCTGGCGGCGCTTGTGCTTATTCATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 1404 GTCCTGGCGGCGCTTGTGCTTATTCATGATGATGATGATGATGATGATGATGATGATGAT 1463
QY 421 ATGCTACTCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTG 480
DB 1464 ATGCTACTCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTG 1523
QY 481 GCGGCGCGAGGCTTATCATCTCTTCACTGCGGCGGCTGAGAAATTCAGCTTGTGA 540
DB 1524 AACACCTCGGAGATTAAGTCTCTTTTCACTGCGGCGGCTGAGAAATTCAGCTTGTGA 1583
QY 541 AACACCAAGCGCGCTGCGCATCAACAGATGCTTGAATGCAATGCTCCCTCCAA 600
DB 1584 AACACCAAGCGCGCTGCGCATCAACAGATGCTTGAATGCAATGCTCCCTCCAA 1643
QY 601 ACTGGGCTCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTG 660
DB 1644 ACTGGGCTCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTG 1703
QY 661 CGGATGGCGAGCTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTG 720
DB 1704 CGGATGGCGAGCTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTG 1763
QY 721 AATGATGCTCAAGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCT 780
DB 1764 AATGATGCTCAAGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCT 1823
QY 781 GGTATGCTGCGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCT 840

```

```

DB 1824 GGTATGCTACCCGCGGCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1883
QY 841 GTGTGTGGGAGCAACCGATGCTTTCGCGCGCTTACGTACGATGAGGTGGAATGAGACG 900
DB 1884 GTGTGTGGGAGCAACCGATGCTTTCGCGCGCTTACGTACGATGAGGTGGAATGAGACG 1943
QY 901 GAGGTGCTGCTTTCACCAACACGCGGCGCGCAAGCGTGTGCGGCGCTTGTGCGGCGCTTGTG 960
DB 1944 GAGGTGCTGCTTTCACCAACACGCGGCGCGCAAGCGTGTGCGGCGCTTGTGCGGCGCTTGTG 1003
QY 961 ATGAAATAGCAACCGGCTTTCACCAAGAGAGTGTGCGGCGCGCGCGCTGCAACATCGGCGGCTC 1020
DB 2004 ATGAAATAGCAACCGGCTTTCACCAAGAGAGTGTGCGGCGCGCGCGCTGCAACATCGGCGGCTC 2063
QY 1021 GGCACACACCTTTGATCTGCCCCAGAGTCTTCCGAGAGATCCGAGGCACTTAC 1080
DB 2064 GGCACATTAACCTTTGATCTGCCCCAGAGTCTTCCGAGAGATCCGAGGCACTTAC 2123
QY 1081 ACCAATAGCGGCTTGTGCGGCGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCTTGTG 1107
DB 2124 ACCAATAGCGGCTTGTGCGGCGCGCTTGTGCGGCGCTTGTGCGGCGCTTGTGCGGCTTGTG 2150

```

RESULT 7  
ID ABK91411 standard; DNA; 9605 BP.  
XX ABK91411;  
XX 15-NOV-2002 (first entry)  
XX  
DE Hepatitis C virus Con 1 isolate DNA.  
XX  
XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
XX Internal ribosome entry site; IRES; NS5A; HCV replication.  
OS Hepatitis C virus.  
XX  
XX Key Location/Qualifiers  
FH CDS 342..9374  
FT /tag= a  
FT /product= "HCV polyprotein"  
FT /note= "The polyprotein consists of the Core, E1, E2, P7,  
FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
XX  
PN MO200259321-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 16-JAN-2002; 2002MO-EP000526.  
XX  
PR 23-JAN-2001; 2001US-0263479P.  
XX  
PA (RICE-) IST RICECHB BIOL MOLECULAR ANGELIETTI.  
PI De Francesco R, Migliaccio G, Paonessa G;  
XX  
XX WPI; 2002-599793/64.  
XX P-PSDB; ABG32451.  
XX  
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX  
PS Claim 9; Page 36-39; 69pp; English.  
XX  
XX The invention relates to nucleic acid molecules comprising altered HCV  
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
CC Internal ribosome entry site (IRES) region coding for one or more NS3,  
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations

CC are detailed in the specification. Also included are (1) an expression vector comprising a nucleotide sequence coding for the altered nucleic acid, which is transcriptionally coupled to an exogenous promoter; (2) a recombinant cell human hepatoma cell comprising the altered nucleic acids; (3) a recombinant cell produced by introducing into a human hepatoma cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus) replicon enhanced cell or which containing a functional HCV replicon; (5) an HCV replicon enhanced cell made in the method; and (6) measuring the ability of a compound to affect HCV activity. The HCV replicons and HCV replicon enhanced cells are useful in studying HCV replication and expression, and HCV and host cell interactions, producing HCV RNA and proteins, and providing a system for measuring the ability of a compound to modulate one or more HCV activities e.g. to discover drugs which may treat HCV mediated diseases such as liver failure, cirrhosis and hepatocellular carcinoma. The present sequence is the HCV replicon Con 1, used as a basis for the adaptive mutations of the invention

XX Sequence 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 U; 0 Other;

Query Match 83.8%; Score 927.8; DB 6; Length 9605;  
Best Local Similarity 89.9%; Pred. No. 5,1e-256;

Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1 TCTTCCCGCTGCTGGTGGTACGCTCTCACTCCACGCTCGGSCCAAGACGCGACATCCCC 60  
DB 1044 TCTTCCCGCTGCTGGTGGTACGCTCTCACTCCACGCTCGGSCCAAGACGCTACGTCGCC 1103  
QY 61 ACTGCGACATATAGACGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
DB 1104 ACTGCGACATATAGACGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1163  
QY 121 ATGTACGTGGGGATCTCTGCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
DB 1164 ATGTACGTGGGGATCTCTGCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1223  
QY 181 CCTGCGACATATAGACGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
DB 1224 CCTGCGCGGACAGACAGATGCAATGCTGCAATATATATATATATATATATATATATATATAT 1283  
QY 241 GGTACCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
DB 1284 GGTACCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343  
QY 301 TCGACGCTACTCCGATATCCCAAGCTGTGCGACATGCTGCGGGGCCCACTGCGGGA 360  
DB 1344 TCGACGCTACTCCGATATCCCAAGCTGTGCGATATGCTGCGGGGCCCACTGCGGGA 1403  
QY 361 GTCTGCGGCGCTTGTCTACTATTCATGCTGCGGGAATGCGGCTAAAGTCTTGTGTG 420  
DB 1404 GTCTGCGGCGCTTGTCTACTATTCATGCTGCGGGAATGCGGCTAAAGTCTTGTGTG 1463  
QY 421 ATGCTACTCTTCCGCGCGCTGACGCGGGAACCTTACACGACAGGGGGGACACAGCGCGC 480  
DB 1464 ATGCTACTCTTCCGCGCGCTGACGCGGGAACCTTATGTGACAGGGGGGACAGATGGCCAA 1523  
QY 481 GCCGCCACGCGGCTTACATCCCTCTTCAACACTGCGCGGCTCAGAAATCCAGCTTGT 540  
DB 1524 AACACCTCGGATATGCTCTCTTTCACCGCGGCTATCCAGAAATCCAGCTTGT 1583  
QY 541 AACACCAAGCGGCTGACATCAACAGAACTGCTTGAATGCAATGAATCCCTTCAA 600  
DB 1584 AACACCAAGCGGCTGACATCAACAGAACTGCTTGAATGCAATGAATCCCTTCAA 1643  
QY 601 ACTGGGTTCTTCCGCGCGCTTGTACACGACAGGTTCAATGCTCGGATGTCTAGAG 660  
DB 1644 ACTGGGTTCTTCCGCGCGCTTGTGTACAGCAAGTTCAATGCTCGGATGTCTAGAG 1703  
QY 661 CGCATGGCGAGTCCGCGCGCTTGAACAGGTTGATGAGGGGTGGGGTCCCATCACTTAT 720  
DB 1704 CGCATGGCGAGTCCGCGCGCTTGAACAGGTTGATGAGGGGTGGGGTCCCATCACTTAC 1763  
QY 721 AATGAGTCCCAAGGCTTGAACAGAGGCTTATGTGCTGACCTAAGCACTCAACCGTGT 780

DB 1764 AATGAGTCCCAAGCTTGAACAGAGGCTTATGTGCTGACTAGCAACCCGCGCTGC 1823  
QY 781 GGTATGTCGCGCGCTTGAAGGTTGTGGCCAGTACTGTTTCACTCAAGCCCTGTT 840  
DB 1824 GGTATGTCGCGCGCGCGCGAGT 1883  
QY 841 GTGGTGGGACGACCGATGTTTGGCGCCCTTACGTAAGATGGGGTGAATGAGACG 900  
DB 1884 GTGGTGGGACGACCGAGCTTGGCGCTTCTTACGTAAGATGGGGGAGATGAGACG 1943  
QY 901 GAGTGTCTCTTCTCAACACAGCGCGCCGACCGGGGAACTGTGCTGTATATAG 960  
DB 1944 GAGTGTCTCTTCTTAAACACGCGCGCGCAAGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2003  
QY 961 ATGATATAGACGCGGTTTCAACAGAGCTGTGGGGGCGCCCGTGCACATCGGGGGGTC 1020  
DB 2004 ATGATATAGACGCTGT 2063  
QY 1021 GGCACACACCTTGT 1080  
DB 2064 GGCATATTAACCTTGAACCTGCGCCCAAGACTGT 2123  
QY 1081 ACCAATGCGGTTTGGGGGCTTGTGTGT 1107  
DB 2124 ACCAATGCTGT 2150

RESULT 8  
ABK91430  
ID ABK91430 standard; DNA; 9605 BP.  
XX  
AC ABK91430;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE Hepatitis C virus Con 1 isolate DNA mutant 7.  
XX  
KM HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KM internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
OS Hepatitis C virus.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 342..9374  
FT /tag= a  
FT /product= "HCV polyprotein"  
FT /note= "The polyprotein consists of the Core, E1, E2, p7,  
FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
FT mutation  
FT replace (6931, C)  
FT /tag= b  
XX  
PN WO200259321-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 16-JAN-2002; 2002WO-EP000526.  
XX  
PR 23-JAN-2001; 2001US-0263479P.  
XX  
PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELLETTI.  
XX  
PI De Francesco R, Migliaccio G, Paonessa G;  
XX  
XX WPI; 2002-599793/64.  
XX  
PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX  
PS Claim 9; Page; 69pp; English.



XX The invention relates to nucleic acid molecules comprising altered HCV  
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acid, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cell made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
 CC mutant of the invention. Note: The present sequence is not shown in the  
 CC specification but was created by the indexer using the HCV sequence  
 CC appearing as ABK91411 and the information in Claim 9

SQ Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 U; 0 Other;

Query Match 83.8%; Score 927.8; DB 6; Length 9605;  
 Best Local Similarity 89.9%; Pred. No. 5.1e-256;  
 Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1 TCCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCGCGGCAAGAGCCGACATCCCC 60  
 DB 1044 TCCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCGCGGCAAGAGCCGACATCCCC 1103

QY 61 ACTGCGACAAATGACAGCGCACTGATGCTGCTGGTGGGCGGCTGCTGCTCGCT 120  
 DB 1104 ACTGCGACAAATGACAGCGCACTGATGCTGCTGGTGGGCGGCTGCTGCTCGCT 1163

QY 121 ATGTAGTGGGGGATCTCGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 DB 1164 ATGTAGTGGGGGATCTCGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1223

QY 181 CCTCGCGACATCAGAGCGTACAGACTGCAATGTTTCAATCTATCCCGGCACTGAT 240  
 DB 1224 CCTCGCGCGACAGAGCGTACAGACTGCAATGTTTCAATCTATCCCGGCACTGAT 1283

QY 241 GGTACCGGATGCTTGGATATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 1284 GGTACCGGATGCTTGGATATGATGATGATGATGATGATGATGATGATGATGAT 1343

QY 301 TGGCAGTACTCCGATTCACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 DB 1344 TGGCAGTACTCCGATTCACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1403

QY 361 GTCCTGGCGGCGCTTCTCTACTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 DB 1404 GTCCTGGCGGCGCTTCTCTACTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1463

QY 421 ATGCTACTTCTGCGCGGCTTGAAGGGAACCTTACACGACAGGCGGGAACACAGCG 480  
 DB 1464 ATGCTACTTCTGCGCGGCTTGAAGGGAACCTTACAGAGGCGGGAACAGATGGCA 1523

QY 481 GCGCGCCAGCGGCTTACATCCCTCTTCAACCTGCGGCGGCTGAGAAATCCAGCTTGA 540  
 DB 1524 AACACCTCGGGAATGATGCTCTCTTTTCAACCGGCGGCTGAGAAATCCAGCTTGA 1583

QY 541 AACACCAAGCGGCTGCGACATCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 DB 1584 AACACCAAGCGGCTGCGACATCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1643

QY 601 ACTGGGTTCTTGGCGCGCTGTTCTACAGCGACAGGTTCAATGCTCGGATCTGAG 660  
 DB 1644 ACTGGGTTCTTGGCGCGCTGTTCTACAGCGACAGGTTCAATGCTGATGATCCAGAG 1703

QY 661 CCGATGGCGAGTGGCGGCGGCAATTGACAGTGCATCAGGGGTGGGATCCACTTAT 720  
 DB 1704 CCGATGGCGAGTGGCGGCGGCAATTGACAGTGCATCAGGGGTGGGATCCACTTAT 1763

QY 721 AATGATGCCAGCGCTTGAACAGAGCGCTTATGCTGCACTACGCACTCAACCGTGT 780  
 DB 1764 AATGATGCCAGCGCTTGAACAGAGCGCTTATGCTGCACTACGCACTCAACCGTGT 1823

QY 781 GGTATGTCGCGCGGCTTGAAGGTGTGAGCCAGTGTACTGTTTCACTCAAGCCGTGT 840  
 DB 1824 GGTATGTCGCGCGGCTTGAAGGTGTGAGCCAGTGTACTGTTTCACTCAAGCCGTGT 1883

QY 841 GTGGTGGGAGCAGCCGATGCTTGGCGGCGGCTTACGTAAGATGGGTGAGAGAG 900  
 DB 1884 GTGGTGGGAGCAGCCGATGCTTGGCGGCGGCTTACGTAAGATGGGTGAGAGAG 1943

QY 901 GACGTGCTCTTCTCAACACAGCGCGCGGCAAGCGGCGGCACTGCTGCTGTACATG 960  
 DB 1944 GACGTGCTCTTCTTAAACACAGCGCGCGGCAAGCGGCGGCACTGCTGCTGTACATG 2003

QY 961 ATGAATAGACCGGCTTCAACAGAGAGTGTGGGGGCGGCGGCAATCGGGGGGCT 1020  
 DB 2004 ATGAATAGACCGGCTTCAACAGAGAGTGTGGGGGCGGCGGCAATCGGGGGGCT 2063

QY 1021 GCGAACACACTTGTATGCTGCGGCGGCAAGCTGCTTCCGGAAGATCCGAGGCACTTAC 1080  
 DB 2064 GCGAACACACTTGTATGCTGCGGCGGCAAGCTGCTTCCGGAAGATCCGAGGCACTTAC 2123

QY 1081 ACCAAATGCGGCTTGGGGGCGGCTTGTG 1107  
 DB 2124 ACCAAATGCGGCTTGGGGGCGGCTTGTG 2150

RESULT 9  
 ABK91428  
 ID ABK91428 standard; DNA; 9605 BP.  
 XX  
 XX ABK91428;  
 AC  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE Hepatitis C virus Con 1 isolate DNA mutant 5.  
 XX  
 KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
 OS  
 OS Hepatitis C virus.  
 OS Synthetic.  
 FH  
 FT Key Location/Qualifiers  
 FT CDS 342..9374  
 FT /\*tag= a  
 FT /product= "HCV polypeptide"  
 FT /note= "The polypeptide consists of the Core, E1, E2, E7,  
 FT mutation NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
 FT replace (6463,A)  
 FT /\*tag= b  
 PN  
 PN MO200259321-A2.  
 PD  
 PD 01-AUG-2002.  
 XX  
 XX 16-JAN-2002; 2002MO-BP000526.  
 PF  
 PF 23-JAN-2001; 2001US-0263479P.  
 PR  
 PR (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.  
 PA  
 PA De Francesco R, Migliaccio G, Paonessa G;  
 PI  
 PI WPI; 2002-599793/64.  
 DR

XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and  
 PT expression.

XX Claim 9; Page; 69pp; English.

XX The invention relates to nucleic acid molecules comprising altered HCV  
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS5a, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acid, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
 CC mutant of the invention. Note: The present sequence is not shown in the  
 CC specification but was created by the indexer using the HCV sequence  
 CC appearing as ABK91411 and the information in Claim 9

XX Sequence 9605 BP; 1909 A; 2884 C; 2733 G; 2079 T; 0 U; 0 Other;

XX Query Match 83.8%; Score 927.8; DB 6; Length 9605;

XX Best Local Similarity 89.9%; Pred. No. 5.1e-256; Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

XX 1 TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCAAGAGCGCAGCATCCCC 60

XX 1044 TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCAAGAGCGTACGCTCC 1103

XX 61 ACTGCACAAATGACAGCGCCAGCTGCAATTTGCTGTTGGGGCGCTGCTTCTCGCT 120

XX 1104 ACTACAGCATGACAGCGCCAGCTGCAATTTGCTGTTGGGGCGCTGCTTCTCGCT 1163

XX 121 ATGTAGTGGGGGATCTCTGCGGATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

XX 1164 ATGTAGTGGGGGATCTCTGCGGATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1223

XX 181 CCTGCGCAGATGACAGCGCTGCAATTTGCTGTTGGGGCGCTGCTTCTCGCT 240

XX 1224 CTCTGCGGCGACAGACAGTACAGAGCTGCAATTTGCTGTTGGGGCGCTGCTTCTCG 1283

XX 241 GGTACCGCATGCTTGGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 300

XX 1284 GGTACCGCATGCTTGGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1343

XX 301 TCGCAGTACTCTCGGATCCCAAGCTGTGTGGAATGTTGGCGGGGCGCACTGGGGA 360

XX 1344 TCGCAGTACTCTCGGATCCCAAGCTGTGTGGAATGTTGGCGGGGCGCACTGGGGA 1403

XX 361 GTCTGTGGCGGCTTGTGCTACTATTCATGTTGGGGAACTGGGCTAAGTCTTGGTGTG 420

XX 1404 GTCTGTGGCGGCTTGTGCTACTATTCATGTTGGGGAACTGGGCTAAGTCTTGGTGTG 1463

XX 421 ATGCTACTCTTGTGGCGGCTTGTGCTACTATTCATGTTGGGGAACTGGGCTAAGTCTT 480

XX 1464 ATGCTACTCTTGTGGCGGCTTGTGCTACTATTCATGTTGGGGAACTGGGCTAAGTCTT 1523

XX 481 GCGGCCACGGGCTTACATCCCTCTTCAACACTGGGCGGGCTCAGAAAATCCAGCTTGT 540

XX 1524 AACACCTCTGGGATTAAGTCTCTTCTTCAACCGGGGCTATCCAGAAAATCCAGCTTGT 1583

QY 541 AACACCAACGCGAGCTGGGACATCAACGAATGCTTGAATCTGCAATGACTCTCCCA 600

DB 1584 AACACCAACGCGAGCTGGGACATCAACGAATGCTTGAATCTGCAATGACTCTCCCA 1643

QY 601 ACTGGGTTCTTCCGCGCTGTTTCAACAGCAGATTCATGCGGCGGATGCTCAGG 660

DB 1644 ACTGGGTTCTTCCGCGCTGTTTCAACAGCAGATTCATGCGGCGGATGCTCAGG 1703

QY 661 CCGATGCGCAGCTGCGCGCGCCCATTTGACAGATTCATGCGGCGGATGCTCAGG 720

DB 1704 CCGATGCGCAGCTGCGCGCGCCCATTTGACAGATTCATGCGGCGGATGCTCAGG 1763

QY 721 AATGATCCACAGCTTGGACCAAGGCGCTTATTTGCTGCACTTACGACCTCAACCGT 780

DB 1764 AATGATCCACAGCTTGGACCAAGGCGCTTATTTGCTGCACTTACGACCTCAACCGT 1823

QY 781 GGTATCTGCGCGCGCTGTTGAGGTGAGTGGCCAGTACTGTTTCACTCAAGCCCTGT 840

DB 1824 GGTATCTGCGCGCGCTGTTGAGGTGAGTGGCCAGTACTGTTTCACTCAAGCCCTGT 1883

QY 841 GTGTGGGACAGCAGATGCTTGGCGCGCCCTTACGTAACAGTGGGAGTGAATGAGAG 900

DB 1884 GTGTGGGACAGCAGATGCTTGGCGCGCCCTTACGTAACAGTGGGAGTGAATGAGAG 1943

QY 901 GACGTGCTGTTCTCAACAAACAGCGCGCCGCAACGAGGCACTGTTGCTGTACATG 960

DB 1944 GACGTGCTGTTCTTAAACAAACAGCGCGCCGCAACGAGGCACTGTTGCTGTACATG 2003

QY 961 ATGATATGACAGCGGTTTCAACAGAGCTGTGGGGGCGCCGTTGCAACATCGGGGGG 1020

DB 2004 ATGATATGACAGCGGTTTCAACAGAGCTGTGGGGGCGCCGTTGCAACATCGGGGGG 2063

QY 1021 GGCACCAACACTTGTATCTGCCCCACAGACTCTTCCGGAACATCCCGAGGCACTTAC 1080

DB 2064 GGCACCAACACTTGTATCTGCCCCACAGACTCTTCCGGAACATCCCGAGGCACTTAC 2123

QY 1081 ACCAAATGCGGTTCCGGGCGCTTGTG 1107

DB 2124 ACCAAATGCGGTTCCGGGCGCTTGTG 2150

RESULT 10

ABK91425

ID ABK91425 standard; DNA; 9605 BP.

XX ABK91425;

XX 15-NOV-2002 (first entry)

XX Hepatitis C virus Con 1 isolate DNA mutant 2.

XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;

XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;

XX internal ribosome entry site; IRES; NS5a; HCV replication; mutant.

XX Hepatitis C virus.

XX Synthetic.

XX

XX

XX

XX

XX

XX

XX

WO200259321-A2.

01-AUG-2002.

16-JAN-2002; 2002WO-EP000526.

Location/Qualifiers  
 /tag= a  
 /product= "HCV polypeptide"  
 /note= "The polypeptide consists of the Core, E1, E2, P7,  
 NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
 replace(3946,A)  
 /tag= b



XX MO200259321-A2.  
XX 01-ANG-2002.  
XX 16-JAN-2002; 2002MO-EP000526.  
XX 23-JAN-2001; 2001US-0263479P.  
XX (RICE-) 1ST RICECHIE BIOL MOLECOLARE ANGELETTI.  
XX De Francesco R, Migliaccio G, Paonessa G;  
XX MPI; 2002-599793/64.  
XX  
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
XX PT ribosome entry site (IRES) region, useful in studying HCV replication and  
XX PT expression.  
XX  
XX Claim 9; Page; 69pp; English.  
XX  
XX The invention relates to nucleic acid molecules comprising altered HCV  
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
XX internal ribosome entry site (IRES) region coding for one or more NS3,  
XX NS5a, or EMCV IRES mutations, respectively. The location of the mutations  
XX are detailed in the specification. Also included are (1) an expression  
XX vector comprising a nucleotide sequence coding for the altered nucleic  
XX acid, which is transcriptionally coupled to an exogenous promoter; (2) a  
XX recombinant cell human hepatoma cell comprising the altered nucleic acids  
XX; (3) a recombinant cell produced by introducing an HCV (hepatitis C virus)  
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
XX replicon enhanced cell or which containing a functional HCV replicon; (5)  
XX an HCV replicon enhanced cells made in the method; and (6) measuring the  
XX ability of a compound to affect HCV activity. The HCV replicons and HCV  
XX replicon enhanced cells are useful in studying HCV replication and  
XX expression, and HCV and host cell interactions, producing HCV RNA and  
XX proteins, and providing a system for measuring the ability of a compound  
XX to modulate one or more HCV activities e.g. to discover drugs which may  
XX treat HCV mediated diseases such as liver failure, cirrhosis and  
XX hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
XX mutant of the invention. Note: The present sequence is not shown in the  
XX specification but was created by the indexer using the HCV sequence  
XX appearing as ABK91411 and the information in Claim 9  
XX  
SQ Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 U; 0 Other;  
Query Match 83.8%; Score 927.8; DB 6; Length 9605;  
Best Local Similarity 89.9%; Pred. No. 5,1e-256;  
Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
QY 1 TCCTCCCGCTGCTGGGAGGCTCACTCCACGCTGCGGCGCAAGAGCGCCACATCCCC 60  
DB 1044 TCTCTCCGCTGCTGGGAGGCTCACTCCACGCTGCGGCGCAAGAGCGCTTACGCTCCC 1103  
QY 61 ACTGCGACATATACGAGCGCAGTCGATTTCTGTTGGGCGGCTGCTGCTCGCT 120  
DB 1104 ACTACGACGATACGAGCGCAGTCGATTTCTGTTGGGCGGCTGCTGCTCGCT 1163  
QY 121 ATGTACGTGGGGGATCTCTGCGGATCTGTTTCTCTGCTCTCACTGTTCACTTCTCG 180  
DB 1164 ATGTACGTGGGAGATCTCTGCGGATCTGTTTCTCTGCTCTCACTGTTCACTTCTCG 1223  
QY 181 CTTGCGCGGACATACAGAGGATGACGATGTCATTTGTTCAATCTCCGCGCACGATCA 240  
DB 1224 CTTGCGCGGACGACAGAGATGACGATGTCATTTGTTCAATCTCCGCGCACGATCA 1283  
QY 241 GGTACCGCGATGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
DB 1284 GGTACCGCGATGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1343  
QY 301 TCGCAGCTACTCGGATCCCAAGCTGTGTGAGACATGATGATGATGATGATGATGATGATG 360  
|||||

DB 1344 TCGCAGTTATCTCCGATCCCAAGCTGTGTGATGATGATGATGATGATGATGATGATG 1403  
QY 361 GTCTGCGCGGCGCTTGGCTACTATTCATGATGATGATGATGATGATGATGATGATGATG 420  
DB 1404 GTCTGCGCGGCGCTTGGCTACTATTCATGATGATGATGATGATGATGATGATGATGATG 1463  
QY 421 ATGCTACTTTTGGCGGCTTGAACGGGAACTTTACACGACAGGGGGACACACGGCGCG 480  
DB 1464 ATGCTACTTTTGGCGGCTTGAACGGGAACTTTACACGAGGGGGACGATGCGCAAA 1523  
QY 481 GCGCGCGACGGCTTACATCCCTTCACACCTGGGCGCGCTCAGAAATTCAGCTTGA 540  
DB 1524 AACACCTTGGGATTTACGTCCTTTTACCCGGGTATCCAGAAATTCAGCTTGA 1583  
QY 541 AACACCAACGAGCTGAGCATCAACAGAACTGCTTGAATGATGATGATGATGATGATGATG 600  
DB 1584 AACACCAACGAGCTGAGCATCAACAGAACTGCTTGAATGATGATGATGATGATGATGATG 1643  
QY 601 ACTGGGTTCTTGGCGGCTTGTTCACACGACAGGTTCAATGCTGCGGATGCTCAGAG 660  
DB 1644 ACTGGGTTCTTGGCGGCTTGTTCACACGACAGGTTCAATGCTGCGGATGCTCAGAG 1703  
QY 661 CGCATGCGCAGCTGCGCGCCCATTTGACAGTTTCAGATGAGGGTGGGCTCCATCACTTAT 720  
DB 1704 CGCATGCGCAGCTGCGCGCCCATTTGACAGTTTCAGATGAGGGTGGGCTCCATCACTTAT 1763  
QY 721 AATGATGCCAGGCTTGGACAGAGGCTTATGCTGACATGACATGACATGACATGACATGAC 780  
DB 1764 AATGATGCCAGGCTTGGACAGAGGCTTATGCTGACATGACATGACATGACATGACATGAC 1823  
QY 781 GGTATGTCGCGCGCTTGGACAGGCTTATGCTGACATGACATGACATGACATGACATGAC 840  
DB 1824 GGTATGTCGCGCGCTTGGACAGGCTTATGCTGACATGACATGACATGACATGACATGAC 1883  
QY 841 GTGATGTCGCGCGCTTGGACAGGCTTATGCTGACATGACATGACATGACATGACATGAC 900  
DB 1884 GTGATGTCGCGCGCTTGGACAGGCTTATGCTGACATGACATGACATGACATGACATGAC 1943  
QY 901 GACGTGCTCTTCTCAACACAGCGCGGCGCCACGAGGAGCACTGCTGCGCTGATCATG 960  
DB 1944 GACGTGCTCTTCTCAACACAGCGCGGCGCCACGAGGAGCACTGCTGCGCTGATCATG 2003  
QY 961 ATGATATGACCGGCTTCAACAGAGCTGAGGAGGCGCGCGCTGCAACATCGGAGGAGTC 1020  
DB 2004 ATGATATGACCGGCTTCAACAGAGCTGAGGAGGCGCGCGCTGCAACATCGGAGGAGTC 2063  
QY 1021 GGCACACACATTTGATGCTGCGCCACGAGACTGTTCCGGAAGATCCGAGGCACTTAC 1080  
DB 2064 GGCACACACATTTGATGCTGCGCCACGAGACTGTTCCGGAAGATCCGAGGCACTTAC 2123  
QY 1081 ACCAAATGCGGTTGCGGAGCTTGGTTG 1107  
DB 2124 ACCAAATGCGGTTGCGGAGCTTGGTTG 2150  
RESULT 12  
ABK91433  
ID ABK91433 standard; DNA; 9605 BP.  
XX  
XX ABK91433;  
AC  
XX  
XX 15-NOV-2002 (first entry)  
DT  
XX  
XX  
DB Hepatitis C virus Con 1 isolate DNA mutant 10.  
XX  
XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
XX internal ribosome entry site; IRES; NS5a; HCV replication; mutant.  
OS Hepatitis C virus.  
OS Synthetic.  
XX  
XX  
FH Key Location/Qualifiers

```

CDS          342..9374
FT            /*tag= a
FT            /product= "HCV polypotein"
FT            /note= "The polypotein consists of the Core, E1, E2, P7,
FT            NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT            mutation
FT            /tag= b
FT            /replace(6953, C)
XX            WO200259321-A2.
XX            01-AUG-2002.
XX            16-JAN-2002; 2002WO-EP000526.
XX            23-JAN-2001; 2001US-0263479P.
XX            (RICE-) IST RICERCHIE BIOL. MOLECOLARE ANGELTETTI.
XX            De Francesco R, Migliaccio G, Paonessa G;
XX            WPI; 2002-599793/64.
XX            New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX            NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX            ribosome entry site (IRES) region, useful in studying HCV replication and
XX            expression.
XX            Claim 9; Page; 69pp; English.
XX            The invention relates to nucleic acid molecules comprising altered HCV
XX            NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX            internal ribosome entry site (IRES) region coding for one or more NS3,
XX            NS5A, or EMCV IRES mutations, respectively. The location of the mutations
XX            are detailed in the specification. Also included are (1) an expression
XX            vector comprising a nucleotide sequence coding for the altered nucleic
XX            acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX            recombinant cell human hepatoma cell comprising the altered nucleic acids
XX            ; (3) a recombinant cell produced by introducing into a human hepatoma
XX            cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX            replicon enhanced cell or which containing a functional HCV replicon; (5)
XX            an HCV replicon enhanced cells made in the method; and (6) measuring the
XX            ability of a compound to affect HCV activity. The HCV replicons and HCV
XX            replicon enhanced cells are useful in studying HCV replication and
XX            expression, and HCV and host cell interactions, producing HCV RNA and
XX            proteins, and providing a system for measuring the ability of a compound
XX            to modulate one or more HCV activities e.g. to discover drugs which may
XX            treat HCV mediated diseases such as liver failure, cirrhosis and
XX            hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
XX            mutant of the invention. Note: The present sequence is not shown in the
XX            specification but was created by the indexer using the HCV sequence
XX            appearing as ABK91411 and the information in Claim 9
XX            Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2079 T; 0 U; 1 Other;

Query Match      83.8%; Score 927.8; DB 6; Length 9605;
Best Local Similarity 89.9%; Pred. No. 5.1e-256;
Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY      1 TCCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCGCGGCAAGAGCCAGCATCCCC 60
DB      1044 TCCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCGCGGCAAGAGCGTACGCTCCC 1103
QY      61 ACTGCGCAATATACAGACGCCAGCTGCATTGTGCTGTTGGGGGCGCGCTTCTGTGCTCGCT 120
DB      1104 ACTATGACGATATACAGACGCCAGCTGCATTGTGCTGTTGGGGGCGCGCTTCTGTGCTCGCT 1163
QY      121 ATGTACGTGGGGGATCTCTGCGGATCTGTTTCTCTGTTCTCTAGACTGTTCACCTTCTCG 180
DB      1164 ATGTACGTGGGGAGATCTCTGCGGATCTGTTTCTCTGTTCTCTAGACTGTTCACCTTCTCG 1223
QY      181 CCTGCGGCAATATACAGACCGCTACAGACCTGCAATGTTAAATCTATCCCGGCAAGTATCA 240
DB      1224 CCTGCGGCGGCAAGACAGTACAGACCTGCAATGTTCTAAATATATATCCCGGCAAGTATCA 1283

```

QY	241	GATCACCAGCATGGCTTGGGAAATATATGATGAACTGGTCACTTACAGCAGCCCTTGTGTA	300
Db	1284	GCTACCCGTATGGCTTGGGAAATATGATGAACTGGTCACTTACAGCAGCCCTTGTGTA	1343
QY	301	TCGCAGCTACTCCGGATCCCAAGCTGTGATGACATGATGGCGGGGGCCACTGGGGA	360
Db	1344	TCGCAGTTACTCCGGATCCCAAGCTGTGATGATGATGGCGGGGGCCCAATTGGGGA	1403
QY	361	GTCTTGGCGGGGCTTGGCTACTATTCCATGGTGGGGAACTGGGCTTAAGTCTTGGTTGTG	420
Db	1404	GTCTTGGCGGGGCTTGGCTACTATTCCATGGTGGGGAACTGGGCTTAAGTCTTGTATGTG	1463
QY	421	ATGCTACTCTTTGGCGGGTGTGACGGGGAACTTACACGACAGGGGGGGAACAAGCGCGC	480
Db	1464	ATGCTACTCTTTGGCGGGTGTGACGGGGGAACTTATGTGACAGGGGGGGAAGATGGCCAA	1523
QY	481	GCCGCCCAAGGGCTTACATCCCTCTTTCACACTTGGGGCGGGCTGAGAAATTCAGCTTGT	540
Db	1524	AAACACCTCGGGATTTACGTCCCTCTTTCACCCGGGTCATCCAGAAATTCAGCTTGT	1583
QY	541	AACACCAACGGCAGCTGGCACAATCAACGAATCTGCTTGAATCTGCATATGATCTCCCTCCA	600
Db	1584	AACACCAACGGCAGCTGGCACAATCAACGAATCTGCTTGAATCTGCATATGATCTCCCTCAAC	1643
QY	601	ACTGGGTTCTTTGGCGGGCTTGTCTTACACGACAGATTCAATGGCTCCGATGCTCAGAG	660
Db	1644	ACTGGGTTCTTTGGCGGGCTTGTCTTACAGTCAACAGATTCAATCTGATGCTCCAGAG	1703
QY	661	CGCATGGCAGCTGCGGCCCATTTGACAGCTTGATCAGGGGTGGGGTCCCATCACTTAT	720
Db	1704	CGCATGGCAGCTGAGGCCCATTCAGCGGTTCCCTCAGGGGTGGGGGCCATCACTTAC	1763
QY	721	AATGATGCCCAAGGCTTGGACCAAGGCCCTTATTTGGGCACTACGCACTTCACCGGT	780
Db	1764	AATGATGTACACAGCTCGGACCAAGGGCTTATTTGGGCACTACGCACTTCACCGGCTGTC	1823
QY	781	GGTATCGGCGCGGGTGTGACAGGTGTGTGGCCAGTGTACTGTTCACCTCCAGGCCCTGTT	840
Db	1824	GGTATCGTACCCGGCGGGCAGGTGTGTGTGCCATGTACTGTCTTCACTCCAGGCCCTGTC	1883
QY	841	GTGGTGGGAGCAGCCGATCTGTTTGGCGGCCCTTACGTACAGATGGGGTGAGATGAGACG	900
Db	1884	GTGGTGGGAGCAGCCGAGCTGGGTGCCCTTACGTACAGTTGGGGGAGAAATGAGACG	1943
QY	901	GACGTGCTGCTTCTTCAACAACAAGCCGGCCGCAAGGGGCAACTGGTTCCGTGTACATGG	960
Db	1944	GACGTGCTGCTTCTTCAACAACAAGCCGGCCGCAAGGGCAACTGGTTGGCTGTACATGG	2003
QY	961	ATGATATAGCAGCGGGTTACCAACAAGCTGTGGGGGCCCCCGGTGCAACATCGGGGGGGTTC	1020
Db	2004	ATGATATAGCAGCTGGGTTTACCAACAAGCTGTGGGGGCCCCCGGTGTACATGTGGGGGATC	2063
QY	1021	GGCAACAACAACCTTGTATCTGCTCCCAAGGACTGCTTCCGGAAGCATCCCGAGGCCACTTAC	1080
Db	2064	GGCAATATAAACCTTGAACCTGGCCCAAGGACTGCTTCCGGAAGCATCCCGAGGCCACTTAC	2123
QY	1081	ACCAATATGCGGTTGGGGGCTTGGTTG 1107	
Db	2124	ACCAATGTTGGTTGGGGGCTTGGTTG 2150	
RESULT 13			
AAD25332			
ID	AAD25332 standard; cDNA, 9605 BP.		
XX	AAD25332;		
XX	AAD25332;		
DT	12-MAR-2002 (first entry)		
XX	Hepatitis C virus (HCV) full-length cDNA mutant #2.		
XX	Hepatitis C virus; HCV; transfection; infection; virus neutralisation;		
KM	Hepatitis C virus; HCV; transfection; infection; virus neutralisation;		



XX Hepatitis C virus Con 1 isolate DNA mutant 4.  
 DE HCV; de; Con 1; adaptive mutation; liver failure; cirrhosis;  
 KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KM internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
 XX Hepatitis C virus.  
 OS Synthetic.  
 XX  
 XX  
 PH Key Location/Qualifiers  
 FT CDS /tag= a  
 FT /product= "HCV polyprotein"  
 FT /note= "The polyprotein consists of the Core, E1, E2, P7,  
 FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
 FT replace(6458..6462,GA)  
 FT mutation /tag= b  
 FT  
 PN MO200259321-A2.  
 XX  
 XX 01-AUG-2002.  
 PD  
 XX 16-JAN-2002; 2002MO-EP000526.  
 PF  
 XX 23-JAN-2001; 2001US-0263479P.  
 PR  
 XX (RICE-) IST RIBERCHE BIOL MOLECULAIRE ANGELETTI.  
 PA De Francesco R, Migliaccio G, Paonessa G;  
 PI  
 XX MPI; 2002-599793/64.  
 DR  
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and  
 PT expression.  
 PT  
 XX  
 PS Claim 9; Page; 69pp; English.

XX The invention relates to nucleic acid molecules comprising altered HCV  
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acid, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC protein, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
 CC mutant of the invention. Note: The present sequence is not shown in the  
 CC specification but was created by the indexer using the HCV sequence  
 CC appearing as ABK91411 and the information in Claim 9  
 CC  
 SQ Sequence 9608 BP; 1913 A; 2883 C; 2733 G; 2079 T; 0 U; 0 Other;

Query Match 83.8%; Score 927.8; DB 6; Length 9608;  
 Best local similarity 89.9%; Pred. No. 5,1e-256;  
 Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1 TCCTCCCGCTGCTGGTAGCGCTCACTCCACGCTGCGGCGCAAGACGCCAGACATCCCC 60  
 DB 1044 TCCTCCCGCTGCTGGTAGCGCTCACTCCACGCTGCGGCGCAAGACGCCAGTACATCCCC 1103

QY 61 ACTGCACAAATACAGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCGCGCT 120  
 DB 1104 ACTACGACGATATAGAGCGCATATGCTGTTGGCTTGGCTGTTGGGGCGGCTGCTTCGCGCT 1163  
 QY 121 ATGTACGTGGGGGATCTTCGCGGATCTGTTTCTGCTGCTCACTGCTTTACCTTCTCG 180  
 DB 1164 ATGTACGTGGGAGATCTTCGCGGATCTGTTTCTGCTGCTGCGCCAGCGTTTCACCTTCTCG 1223  
 QY 181 CCTGCGGACATAGACGGGATGAGGACGTGCAATTTGTTCAATCATTCGCCGCGCAGTATCA 240  
 DB 1224 CCTGCGGACGAGACGAGACGAGACGTGCAATTTGTTCAATCATTCGCCGCGCAGTATCA 1283  
 QY 241 GGTACCGGATGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 1284 GGTACCGGATGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1343  
 QY 301 TCGCAGCTACTCGGATTCGACCAAGCTGTGCTGAGCATGCTGCGGGGCGCCACTGCGGGA 360  
 DB 1344 TCGCAGCTACTCGGATTCGACCAAGCTGTGCTGAGCATGCTGCGGGGCGCCACTGCGGGA 1403  
 QY 361 GTCCTGGCGGGGCTTGGCTACTATTCATGATGAGGGAACCTGGGCTTAAGGCTTGGTTGG 420  
 DB 1404 GTCCTAGCGGGGCTTGGCTACTATTCATGATGAGGGAACCTGGGCTTAAGGCTTGGTTGG 1463  
 QY 421 ATGCTACTCTTTGCGCGGCTTGAAGGGAACCTTACACGACAGGGGGGACACAGGCGCGC 480  
 DB 1464 ATGCTACTCTTTGCGCGGCTTGAAGGGAACCTTACACGAGGGGGGAGATGAGGCGCGCAA 1523  
 QY 481 GCCGCCACGGGCTTACATTCCTCTTTCACACCTGCGGCGGCTGAGAAAATCCAGCTTTGTA 540  
 DB 1524 AACACCTCGGGATATGATGCTTCCTTTTACACCGGGGCTCATCCAGAAAATCCAGCTTTGTA 1583  
 QY 541 AACACCAAGGCGAGCTGGGACATCAACAGAACTGCTTGAACCTGCAATGATCTCCCTCCCA 600  
 DB 1584 AACACCAAGGCGAGCTGGGACATCAACAGAACTGCTTGAACCTGCAATGATCTCCCTCCCA 1643  
 QY 601 ACTGGGTTCTTTCGCGGCTTGTCTACACGACAGGTTCAATGCTGCGGATCTGAGAG 660  
 DB 1644 ACTGGGTTCTTTCGCGGCTTGTCTACACGACAGGTTCAATGCTGCGGATCTGAGATCCAGAG 1703  
 QY 661 CGCATGGCAGCTGCGGCGCCCATTTGACAGTTGATCAGGGGTGGGCTCCCATCACTTAT 720  
 DB 1704 CGCATGGCAGCTGCGGCGCCCATTTGACAGGTTGCTCAGGGGTGGGCGCCCATCACTTAT 1763  
 QY 721 AATGATGCCAGGCTTGGACCAAGAGCCCTATGCTGGGACCTACGACCACTCAACCTGT 780  
 DB 1764 AATGATGCCAGGCTTGGACCAAGAGCCCTATGCTGGGACCTACGACCACTCAACCTGT 1823  
 QY 781 GGTATGTCGCGCGCTTGGACGAGTGTGGCCGAGTACTGTTTCACTCCAGCCCTGTT 840  
 DB 1824 GGTATGTCGCGCGCTTGGACGAGTGTGGCCGAGTACTGTTTCACTCCAGCCCTGTT 1883  
 QY 841 GTGGTGGGAGCAGCCGATGTTTGGCGGCCCTTACGTACAGATGGGGTGAATGAGAGC 900  
 DB 1884 GTGGTGGGAGCAGCCGATGTTTGGCGGCCCTTACGTACAGTGGGGGGAATGAGAGC 1943  
 QY 901 GAGTGTCTCTTTCACCAACAGAGCGCGGCCACAGGGGCACTGTGCTGCTGTATATG 960  
 DB 1944 GAGTGTCTCTTTCACCAACAGAGCGCGGCCACAGGGGCACTGTGCTGCTGTATATG 2003  
 QY 961 ATGAATGACCGGGGTTTCAAGAGAGTGTGGGGGGCCCCCGGACCAATCGGGGGGGGTC 1020  
 DB 2004 ATGAATGACCGGGGTTTCAAGAGAGTGTGGGGGGCCCCCGGACCAATCGGGGGGGGTC 2063  
 QY 1021 GGGCAACACATTTGATCTGCGCCACGAGCTGTTCCGGAAGATCCGAGGCACTTAC 1080  
 DB 2064 GGGCAATTAACCTTGAACCTGCGCCACGAGCTGTTCCGGAAGACCCGAGGCACTTAC 2123  
 QY 1081 ACCAATGCGGTTTCGGGGCTTGGTTG 1107  
 DB 2124 ACCAATGCGGTTTCGGGGCTTGGTTG 2150







OY 1081 ACCAATGCGGTTGCGGGCCTTG 1107  
||| |  
Db 3581 ACCAAGTGTGTTGCGGGCCTTG 3607

Search completed: February 21, 2005, 06:24:51  
Job time : 709 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 05:33:47 : Search time 228 Seconds  
(without alignments)  
7944.554 Million cell updates/sec

Title: US-09-664-363-19

Perfect score: 1107  
Sequence: 1 TCCTCCCGCTGCTGGTAC.....GGGTTGGGCTGTTG 1107

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/6CTUS\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1107	100.0	1107	3	US-08-191-160-19 Sequence 19, Appl
2	1107	100.0	2116	3	US-08-191-160-21 Sequence 21, Appl
3	927.8	83.8	11076	4	US-09-539-601-1 Sequence 1, Appl
4	927.8	83.8	11076	4	US-09-539-601-19 Sequence 19, Appl
5	927.8	83.8	11076	4	US-09-539-601-25 Sequence 25, Appl
6	927.8	83.8	11076	4	US-09-539-601-31 Sequence 31, Appl
7	909.8	82.2	9595	3	US-09-014-416-4 Sequence 4, Appl
8	909.8	82.2	9595	3	US-09-014-416-6 Sequence 6, Appl
9	893.2	80.7	9472	4	US-08-150-204E-96 Sequence 96, Appl
10	891	80.5	9413	4	US-09-827-688-6 Sequence 6, Appl
11	886.2	80.1	6039	1	US-08-324-977-11 Sequence 11, Appl
12	886.2	80.1	6039	2	US-08-384-616-11 Sequence 11, Appl
13	886.2	80.1	6039	2	US-08-304-686A-11 Sequence 11, Appl
14	886.2	80.1	6039	3	US-09-315-850-11 Sequence 11, Appl
15	886.2	80.1	9030	1	US-08-324-977-13 Sequence 13, Appl
16	886.2	80.1	9030	2	US-08-384-616-13 Sequence 13, Appl
17	886.2	80.1	9030	2	US-08-304-686A-13 Sequence 13, Appl
18	886.2	80.1	9030	3	US-09-315-850-13 Sequence 13, Appl
19	886.2	80.1	9416	1	US-08-324-977-1 Sequence 1, Appl
20	886.2	80.1	9416	2	US-08-384-616-1 Sequence 1, Appl
21	886.2	80.1	9416	2	US-08-904-686A-1 Sequence 1, Appl
22	886.2	80.1	9416	3	US-09-315-850-1 Sequence 1, Appl
23	886.2	80.1	9416	3	US-08-823-895A-27 Sequence 27, Appl
24	883	79.8	2082	3	US-08-612-973-47 Sequence 47, Appl
25	883	79.8	2082	3	US-08-927-597-47 Sequence 47, Appl
26	883	79.8	2433	3	US-08-612-973-49 Sequence 49, Appl
27	883	79.8	2433	3	US-08-927-597-49 Sequence 49, Appl

28	689.4	62.3	1539	2	US-08-470-426B-17 Sequence 17, Appl
29	689.4	62.3	1863	2	US-08-470-426B-14 Sequence 14, Appl
30	676.6	61.1	1620	4	US-09-194-949A-10 Sequence 10, Appl
31	664.8	60.1	1476	3	US-08-612-973-35 Sequence 35, Appl
32	664.8	60.1	1476	3	US-08-927-597-35 Sequence 35, Appl
33	642.6	58.0	9646	3	US-08-811-566-1 Sequence 1, Appl
34	642.6	58.0	9646	3	US-09-034-756-1 Sequence 1, Appl
35	642.6	58.0	12980	3	US-08-811-566-5 Sequence 5, Appl
36	642.6	58.0	12980	3	US-09-034-756-5 Sequence 5, Appl
37	639.4	57.8	1207	1	US-08-460-806-1 Sequence 1, Appl
38	639.4	57.8	1207	1	US-08-325-630-1 Sequence 1, Appl
39	639.4	57.8	9599	3	US-09-014-416-2 Sequence 2, Appl
40	636.2	57.5	9365	4	US-09-827-688-7 Sequence 7, Appl
41	634.6	57.3	9401	2	US-08-432-693-1 Sequence 1, Appl
42	634.6	57.3	9416	3	US-08-811-566-19 Sequence 19, Appl
43	634.6	57.3	9416	3	US-09-034-756-19 Sequence 19, Appl
44	633	57.2	1207	1	US-08-460-806-3 Sequence 3, Appl
45	633	57.2	1207	1	US-08-325-630-3 Sequence 3, Appl

#### ALIGNMENTS

RESULT 1  
US-08-191-160-19  
; Sequence 19, Application US/08191160  
; Patent No. 6210675  
; GENERAL INFORMATION:  
; APPLICANT: Highfield, Peter Edmund  
; APPLICANT: Rodgers, Brian Colin  
; APPLICANT: Tedder, Richard Seton  
; APPLICANT: Barbara, John Anthony James  
; TITLE OF INVENTION: Viral Agent  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSER: Rothwell, Figg, Ernst & Kurz  
; STREET: 1700 K Street  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage  
; OPERATING SYSTEM: IBM AT compatible  
; SOFTWARE: Wordperfect 5.0 (DOS text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/191,160  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/628,516  
; FILING DATE: 17 DEC 1990  
; APPLICATION NUMBER: UK 89 28 562.1  
; FILING DATE: 18 DEC 1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 414.0  
; FILING DATE: 27 FEB 1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 814.1  
; FILING DATE: 03 MAR 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: E. Anthony Figg  
; REGISTRATION NUMBER: 27,195  
; REFERENCE/DOCKET NUMBER: 1645-103A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 833-5740  
; TELEFAX: (202) 833-5744  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1107 base pairs  
; TYPE: nucleotide with corresponding protein  
; STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: cDNA to genomic RNA  
ORIGINAL SOURCE:  
ORGANISM: human; serum infectious for PT-NANBH  
IMMEDIATE SOURCE: clone 136/155  
FEATURE:  
LOCATION: from 1 to 1107 bp portion of the PT-NANBH  
LOCATION: polyprotein  
OTHER INFORMATION: probably encodes viral structural  
OTHER INFORMATION: proteins  
US-08-191-160-19

Query Match 100.0%; Score 1107; DB 3; Length 1107;  
Best Local Similarity 100.0%; Pred. No. 1.6e-305;  
Matches 1107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTCCCGCTGCTGGGATGCGCTCACTCCACGCTGGCGGCAAGACGCGACATCCCC 60  
DB 1 TCCTCCCGCTGCTGGGATGCGCTCACTCCACGCTGGCGGCAAGACGCGACATCCCC 60  
QY 61 ACTGCGACATAGAGCGCCAGTCGATTTGCTGGGGGCGGCTTCTGCTCGCT 120  
DB 61 ACTGCGACATAGAGCGCCAGTCGATTTGCTGGGGGCGGCTTCTGCTCGCT 120  
QY 121 ATGTAGTGGGGATCTGCGGATCTGTTTCTGCTCTGAGCTGTTCACCTTCTG 180  
DB 121 ATGTAGTGGGGATCTGCGGATCTGTTTCTGCTCTGAGCTGTTCACCTTCTG 180  
QY 181 CCTGCGGACATAGAGCGGTACAGACTGCAATTTGTTCAATCTATCCCGGCACTATCA 240  
DB 181 CCTGCGGACATAGAGCGGTACAGACTGCAATTTGTTCAATCTATCCCGGCACTATCA 240  
QY 241 GGTACCGCATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 241 GGTACCGCATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
QY 301 TCGCAGCTACTCCGATCCCAAGCTGTGAGCACTGAGCGGCGGCGCACTGGGGA 360  
DB 301 TCGCAGCTACTCCGATCCCAAGCTGTGAGCACTGAGCGGCGGCGCACTGGGGA 360  
QY 361 GTCTGAGCGGCGCTTGGCTTATCTATCTGATGAGGGAAGCTGGGCTAAAGTCTTGTG 420  
DB 361 GTCTGAGCGGCGCTTGGCTTATCTATCTGATGAGGGAAGCTGGGCTAAAGTCTTGTG 420  
QY 421 ATGCTACTCTTTCGCGGCTTGAAGGGAAGCTTAAACGACAGGAGGGAACACGCGCGC 480  
DB 421 ATGCTACTCTTTCGCGGCTTGAAGGGAAGCTTAAACGACAGGAGGGAACACGCGCGC 480  
QY 481 GCCGCCACGAGGCTTACATCCCTTTCACACTGAGCGGCGGCTGAGAAATCAGCTTGT 540  
DB 481 GCCGCCACGAGGCTTACATCCCTTTCACACTGAGCGGCGGCTGAGAAATCAGCTTGT 540  
QY 541 AACAACGAGGCTGAGCTGACATCAACAGAACTGCTTGAATGCAATGATCCCTCCAA 600  
DB 541 AACAACGAGGCTGAGCTGACATCAACAGAACTGCTTGAATGCAATGATCCCTCCAA 600  
QY 601 ACTGGGTTCTTTCGCGGCTTGAAGGGAAGCTTAAAGCTGAGGATGCTGAGAG 660  
DB 601 ACTGGGTTCTTTCGCGGCTTGAAGGGAAGCTTAAAGCTGAGGATGCTGAGAG 660  
QY 661 CGCATGGCAGCTGCGGCTTGAAGGGAAGCTTAAAGCTGAGGATGCTGAGAG 720  
DB 661 CGCATGGCAGCTGCGGCTTGAAGGGAAGCTTAAAGCTGAGGATGCTGAGAG 720  
QY 721 AATGAGTCCACGAGCTTGAAGGGAAGCTTAAAGCTGAGGATGCTGAGAG 780  
DB 721 AATGAGTCCACGAGCTTGAAGGGAAGCTTAAAGCTGAGGATGCTGAGAG 780  
QY 781 GGTATGCTGCGGCTTGAAGGGAAGCTTAAAGCTGAGGATGCTGAGAG 840  
DB 781 GGTATGCTGCGGCTTGAAGGGAAGCTTAAAGCTGAGGATGCTGAGAG 840

QY 841 GTGTGGGAGCAGCCGATGCTTTCGCGGCGCTTACCTACATGATGGGATGAGAGC 900  
DB 841 GTGTGGGAGCAGCCGATGCTTTCGCGGCGCTTACCTACATGATGGGATGAGAGC 900  
QY 901 GACGTGCTTCTTCAACACAGCGGCGGCGGCAAGCTGATGGGCTGTACATG 960  
DB 901 GACGTGCTTCTTCAACACAGCGGCGGCGGCAAGCTGATGGGCTGTACATG 960  
QY 961 ATGATAGCAGCGGCTTCAACACAGCGGCGGCGGCAAGCTGATGGGCTGTACATG 1020  
DB 961 ATGATAGCAGCGGCTTCAACACAGCGGCGGCGGCAAGCTGATGGGCTGTACATG 1020  
QY 1021 GCGAACACACTTGTATGCTGCGGCGGCGGCAAGCTGATGGGCTGTACATG 1080  
DB 1021 GCGAACACACTTGTATGCTGCGGCGGCGGCAAGCTGATGGGCTGTACATG 1080  
QY 1081 ACCAAATGCGGTTCGCGGCGGCTTGTG 1107  
DB 1081 ACCAAATGCGGTTCGCGGCGGCTTGTG 1107

RESULT 2  
US-08-191-160-21  
Sequence 21, Application US/08191160  
Patent No. 6210675  
GENERAL INFORMATION:  
APPLICANT: Highfield, Peter Edmund  
APPLICANT: Rodgers, Brian Colin  
APPLICANT: Tedders, Richard Seton  
APPLICANT: Barbara, John Anthony James  
TITLE OF INVENTION: Viral Agent  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz  
STREET: 1700 K Street  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage  
COMPUTER: IBM AT compatible  
OPERATING SYSTEM: MS-DOS V3.2  
SOFTWARE: Wordperfect 5.0 (DOS text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/191,160  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/628,516  
FILING DATE: 17 DEC 1990  
APPLICATION NUMBER: UK 89 28 562.1  
FILING DATE: 18 DEC 1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 90 04 414.0  
FILING DATE: 27 FEB 1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 90 04 814.1  
FILING DATE: 03 MAR 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: E. Anthony Figg  
REGISTRATION NUMBER: 27,195  
REFERENCE/DOCKET NUMBER: 1645-103A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 833-5740  
TELEFAX: (202) 833-5744  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 216 base pairs  
TYPE: nucleotide with corresponding protein  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to genomic RNA

```
ORIGINAL SOURCE:
ORGANISM: human; serum infectious for PT-NANBH
IMMEDIATE SOURCE:
LIBRARY: cDNA clones from 5' end of the genome
FEATURE:
LOCATION: from 308 to 2116 bp start of the PT-NANBH
LOCATION: polyprotein
OTHER INFORMATION: viral structural and non-structural
OTHER INFORMATION: proteins
US-08-191-160-21

Query Match      100.0%; Score 1107; DB 3; Length 2116;
Best Local Similarity 100.0%; Pred. No. 2e-305;
Matches 1107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTCCCGCTGCTGGGTAGGCTCACTCCCAAGCTGGCGCCCAAGACGACATATCCC 60
DB 1010 TCCTCCCGCTGCTGGGTAGGCTCACTCCCAAGCTGGCGCCCAAGACGACATATCCC 1069
QY 61 ACTGCACAATACGACGCGCAGTCGATTTGCTGTTGGGGGCGCTGCTGCTCGCT 120
DB 1070 ACTGCACAATACGACGCGCAGTCGATTTGCTGTTGGGGGCGCTGCTGCTCGCT 1129
QY 121 ATGTACGTGGGGATCTGCGGATCTGTTTCTGCTCTGCTGCTGCTGCTGCTGCT 180
DB 1130 ATGTACGTGGGGATCTGCGGATCTGTTTCTGCTCTGCTGCTGCTGCTGCTGCTG 1189
QY 181 CTTGCGCGACATACAGCGGTACGACTGCAATTTGTCATCTATCCCGCCACGATCA 240
DB 1190 CTTGCGCGACATACAGCGGTACGACTGCAATTTGTCATCTATCCCGCCACGATCA 1249
QY 241 GGTACACGAGTGGCTGGGATATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 1250 GGTACACGAGTGGCTGGGATATGATGATGATGATGATGATGATGATGATGATGAT 1309
QY 301 TCGACACTACTCCGATTCACCAAGCTGTCTGAGACATGCTGCGGGGCGCCCACTG 360
DB 1310 TCGACACTACTCCGATTCACCAAGCTGTCTGAGACATGCTGCGGGGCGCCCACTG 1369
QY 361 GTCTGCGGGGCTTGTGCTTATTTCCATGTTGGGGAATCGGCTAAAGTCTTGTG 420
DB 1370 GTCTGCGGGGCTTGTGCTTATTTCCATGTTGGGGAATCGGCTAAAGTCTTGTG 1429
QY 421 ATGCTACTCTTTGCGGGGCTTGAAGGGGAACCTTACACGACAGGGGGGACACAG 480
DB 1430 ATGCTACTCTTTGCGGGGCTTGAAGGGGAACCTTACACGACAGGGGGGACACAG 1489
QY 481 GCGGCCACGAGGCTTACATCCCTTTCACACTGCGCGGCTGAGAAATCCAGTTGTA 540
DB 1490 GCGGCCACGAGGCTTACATCCCTTTCACACTGCGCGGCTGAGAAATCCAGTTGTA 1549
QY 541 AACACCAACGCGAGCTGGACATCAACAGAACTGCTTGAATGCAATGACTCCCTCA 600
DB 1550 AACACCAACGCGAGCTGGACATCAACAGAACTGCTTGAATGCAATGACTCCCTCA 1609
QY 601 ACTGGGTTCTTGGCGCGCTTCTTCAACGACAGGTTCAATGCGTCCGATCTCAAG 660
DB 1610 ACTGGGTTCTTGGCGCGCTTCTTCAACGACAGGTTCAATGCGTCCGATCTCAAG 1669
QY 661 CGCATGGCAGTGGCGCCCATTTGACAGATTGATCAGGGGTGGGGTCCCATCACTAT 720
DB 1670 CGCATGGCAGTGGCGCCCATTTGACAGATTGATCAGGGGTGGGGTCCCATCACTAT 1729
QY 721 AATGATCCACGCGTTTGAACAGAGCCCTTATTTGCTGCGACATACGACCTCAAC 780
DB 1730 AATGATCCACGCGTTTGAACAGAGCCCTTATTTGCTGCGACATACGACCTCAAC 1789
QY 781 GGTATGTGCGCGGTTTGAAGGTGTGAGCGCAAGTGTACTGTTCACTCCAAACCGT 840
DB 1790 GGTATGTGCGCGGTTTGAAGGTGTGAGCGCAAGTGTACTGTTCACTCCAAACCGT 1849
QY 841 GTGGTGGGACGACCGATCTTTGCGCGCCCTTACGATACAGATGGGGTGAAATGACG 900
```

```
DB 1850 GTGGTGGGACGACCGATCTTTGCGCGCCCTTACGATACGATGGGGTGAAATGACG 1909
QY 901 GACGTGCTGTTTCTCAACAACAGCGGGCGCCACGAGGAGCACTGTTGCGCTGTACATG 960
DB 1910 GACGTGCTGTTTCTCAACAACAGCGGGCGCCACGAGGAGCACTGTTGCGCTGTACATG 1969
QY 961 ATGATATGACCGCGGTTTCAACAAGAGTGGGGGCCCCCGGTGCAATCGGGGGGCTC 1020
DB 1970 ATGATATGACCGCGGTTTCAACAAGAGTGGGGGCCCCCGGTGCAATCGGGGGGCTC 2029
QY 1021 GGCACACACATTTGATCTGCCCCACGACTGCTTCCGAGACATCCGAGGCACTTAC 1080
DB 2030 GGCACACACATTTGATCTGCCCCACGACTGCTTCCGAGACATCCGAGGCACTTAC 2089
QY 1081 ACCAATGCGGTTTGGGGGCTTGTGTTG 1107
DB 2090 ACCAATGCGGTTTGGGGGCTTGTGTTG 2116

RESULT 3
US-09-539-601-1
; Sequence 1, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartschlagel, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/wt
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polyprotein from core to
; OTHER INFORMATION: nonstructural protein NS5B; parental sequence
; OTHER INFORMATION: without cell culture-adaptive mutations
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
; US-09-539-601-1

Query Match      83.8%; Score 927.8; DB 4; Length 11076;
Best Local Similarity 89.9%; Pred. No. 4.6e-254;
Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

121 ATGTAGTGGGGGATCTCTGCGGATCTGTTTCTCTGCTCTGAGTGTTCACCTTCTCG 180  
2635 ATGTAGTGGGGGATCTCTGCGGATCTGTTTCTCTGCTCTGAGTGTTCACCTTCTCG 2694  
181 CCTGCGGACATCAGACGGTACAGACCTGCAATGTTCAATCTATCCCGGACGATATCA 240  
2695 CCTGCGGACATCAGACGGTACAGACCTGCAATGTTCAATCTATCCCGGACGATATCA 2754  
241 GGTACACGATGCTGTTGGGATATGATGATGAACTGTGACCTTACAGAGCCCTAGTGGTA 300  
2755 GGTACACGATGCTGTTGGGATATGATGATGAACTGTGACCTTACAGAGCCCTAGTGGTA 2814  
301 TCGACACTACTCCGATATCCACAGAGTGTCTGTGAGATGTTGGGCGGCGCCACTGCGGA 360  
2815 TCGACACTACTCCGATATCCACAGAGTGTCTGTGAGATGTTGGGCGGCGCCACTGCGGA 2874  
361 GTCTGAGCGGGCTTGTGCTACTATTCATATGTTGGGGAATGGGCTTAAAGTCTTGTGTG 420  
2875 GTCTGAGCGGGCTTGTGCTACTATTCATATGTTGGGGAATGGGCTTAAAGTCTTGTGTG 2934  
421 ATGCTACTCTTTCGCGCGGTGACGGGGAACCTTACACGACAGGGGGGACACAGGCGCG 480  
2935 ATGCTACTCTTTCGCGCGGTGACGGGGAACCTTATGTGACAGGGGGGACAGATGCGCATA 2994  
481 GCCGCCACAGCGGCTTACATCCCTCTTCAACACTGCGCGGCTGAGAAATCCAGCTTGTGA 540  
2995 AACACCTCGGGATTTAGTCCCTTTTTCACCGCGGCTACCCAGAAATCCAGCTTGTGA 3054  
541 AACACCAACGGGACGTGGACATCAACAGAACTGCTTGAATGCAATGATGCTCTCCCA 600  
3055 AACACCAACGGGACGTGGACATCAACAGAACTGCTTGAATGCAATGATGCTCTCCCA 3114  
601 ACTGGGTTCTTTCGCGCGGTCTCTCAAGCAGAGTGTCAATGGTCCGATGCTAGAG 660  
3115 ACTGGGTTCTTTCGCGCGGTCTCTCAAGCAGAGTGTCAATGGTCCGATGCTAGAG 3174  
661 CGCATGGGACGTCGCGCCCATATGACCAAGTGTGATGAGGGTGGGGTCCCATCACTTAT 720  
3175 CGCATGGGACGTCGCGCCCATATGACCAAGTGTGATGAGGGTGGGGTCCCATCACTTAT 3234  
721 AATGAGTCCACGCGCTTGAACAGAGCCCTTATGCTGAGCACTTACGCACTTCAACGCTGT 780  
3235 AATGAGTCCACGCGCTTGAACAGAGCCCTTATGCTGAGCACTTACGCACTTCAACGCTGT 3294  
781 GGTATGTCGCGCGGTGAGGTGTGAGGCGGCAAGTGTGATGCTTCAACGCGCTGT 840  
3295 GGTATGTCGCGCGGTGAGGTGTGAGGCGGCAAGTGTGATGCTTCAACGCGCGTGT 3354  
841 GTGGTGGGACGACCGATCGTTTCGCGCCCTTACGATACAGATGGGGTGAAGATGAGACG 900  
3355 GTGGTGGGACGACCGATCGTTTCGCGCCCTTACGATACAGATGGGGTGAAGATGAGACG 3414  
901 GACGTGCTGCTTCAACACAGCGCGCGCCCAACGGGGCAATGTGTTCCGCTGATACG 960  
3415 GACGTGCTGCTTCAACACAGCGCGCGCCCAACGGGGCAATGTGTTCCGCTGATACG 3474  
961 ATGAAATGCAACCGGTTCAACAAGACGTGAGGGGGCGCCCGGTGCAACATCGGGGGGTC 1020  
3475 ATGAAATGCAACCGGTTCAACAAGACGTGAGGGGGCGCCCGGTGCAACATCGGGGGGTC 3534  
1021 GGCACAACTTGTATCTCCCAACGACCTGCTTCCGAAAGATCCCGAGGCACTTAC 1080  
3535 GGCACAACTTGTATCTCCCAACGACCTGCTTCCGAAAGATCCCGAGGCACTTAC 3594  
1081 ACCAAATGCGGTTTCGGGGCTTGTG 1107  
3595 ACCAAATGCGGTTTCGGGGCTTGTG 3621

RESULT 4  
US-09-539-601-19  
; Sequence 19, Application US/09539601C

Patent No. 6630343  
GENERAL INFORMATION:  
APPLICANT: Bartschlagel, Ralf FW  
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
FILE REFERENCE: all sequences  
CURRENT APPLICATION NUMBER: US/09/539,601C  
EARLIER FILING DATE: 2001-08-30  
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 19  
LENGTH: 11076  
TYPE: DNA  
ORGANISM: Hepatitis C virus  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: (1)..(341)  
OTHER INFORMATION: construct I389/Core-3'/9-13F  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (342)..(1193)  
OTHER INFORMATION: hepatitis C virus core - neomycin  
OTHER INFORMATION: phosphotransferase fusion protein  
FEATURE:  
NAME/KEY: RBS  
LOCATION: (1202)..(1812)  
OTHER INFORMATION: internal ribosome entry site from  
OTHER INFORMATION: encephalomyocarditis virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1813)..(10845)  
OTHER INFORMATION: hepatitis C virus polyprotein from core to  
OTHER INFORMATION: nonstructural protein NS5B; carries cell  
OTHER INFORMATION: culture-adaptive mutations from clone 9-13F  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: (10846)..(11076)  
US-09-539-601-19

Query Match 83.8%; Score 927.8; DB 4; Length 11076;  
Best Local Similarity 89.9%; Pred. No. 4,6e-254;  
Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

1 TCCTCCCGCTGCTGGGTAGCGCTCACTCCCAAGCTGCGCGCAAGACCCAGATCCCC 60  
2515 TCCTCCCGCTGCTGGGTAGCGCTCACTCCCAAGCTGCGCGCAAGACCCAGATCCCC 2574  
61 ACTGCAATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTGTCCGCT 120  
2575 ACTGCAATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTGTCCGCT 2634  
121 ATGTAGTGGGGGATCTCTGCGGATCTGTTTCTCTGCTCTGAGTGTTCACCTTCTCG 180  
2635 ATGTAGTGGGGGATCTCTGCGGATCTGTTTCTCTGCTCTGAGTGTTCACCTTCTCG 2694  
181 CCTGCGGACATCAGACGGTACAGACCTGCAATGTTCAATCTATCCCGGACGATATCA 240  
2695 CCTGCGGACATCAGACGGTACAGACCTGCAATGTTCAATCTATCCCGGACGATATCA 2754  
241 GGTACACGATGCTGTTGGGATATGATGATGAACTGTGACCTTACAGAGCCCTAGTGGTA 300  
2755 GGTACACGATGCTGTTGGGATATGATGATGAACTGTGACCTTACAGAGCCCTAGTGGTA 2814  
301 TCGACACTACTCCGATATCCACAGAGTGTCTGTGAGATGTTGGGCGGCGCCACTGCGGA 360  
2815 TCGACACTACTCCGATATCCACAGAGTGTCTGTGAGATGTTGGGCGGCGCCACTGCGGA 2874  
361 GTCTGAGCGGGCTTGTGCTACTATTCATATGTTGGGGAATGGGCTTAAAGTCTTGTGTG 420  
2875 GTCTGAGCGGGCTTGTGCTACTATTCATATGTTGGGGAATGGGCTTAAAGTCTTGTGTG 2934  
421 ATGCTACTCTTTCGCGCGGTGACGGGGAACCTTACACGACAGGGGGGACACAGGCGCG 480

```

Db      2935 ATCTACTCTTTCCCGCGCTTGACGGGGAACTATGTGACAGGGGGAGATGGCCAA 2994
Qy      481 GCCGCCAAGGGGCTTACATCTCTTTCAACCTGGGCGGCTCAAAAATCCAGCTTGA 540
Db      2995 AACACCTCGGGGATTCGTCTCTTTTCAACCGGGGTCATCCAGAAATCCAGCTTGA 3054
Qy      541 AACACCAAGGGGAGCTGGGACATCAACAGAACTGCTGAATGCAATGATCTCCCTCA 600
Db      3055 AACACCAAGGGGAGCTGGGACATCAACAGAACTGCTGAATGCAATGATCTCCCTCA 3114
Qy      601 ACTGGTCTCTTCCCGCGCTGTTCTACACGACAGAGTTCAATGCTCCGATCTCAGAG 660
Db      3115 ACTGGTCTCTTCCCGCGCTGTTCTACAGTGAAGTTCATCTGATGATCCAGAG 3174
Qy      661 CGCATGCGCAGCTGCGCGCCCATTTGACAGAACTGATCAGGGGTGGGGTCCCATCTTAT 720
Db      3175 CGCATGCGCAGCTGCGCGCCCATTTGACAGAACTGATCAGGGGTGGGGTCCCATCTTAT 3234
Qy      721 AATGATCCCAAGGCTTGGACAGAGGCTTATGCTGGGACCTACGACCTCAACCGTGT 780
Db      3235 AATGATCCCAAGGCTTGGACAGAGGCTTATGCTGGGACCTACGACCTCAACCGTGT 3294
Qy      781 GGTATGTCGCGCGCTGTTGAGTGTGTGCGCCAGTGTACTGTTCACTCCAGCCCTGT 840
Db      3295 GGTATGTCGCGCGCGCTGTTGAGTGTGTGCGCCAGTGTACTGTTCACTCCAGCCCTGT 3354
Qy      841 GTGGTGGGAGCAGCCGATGTTTCGGGCGCCCTTACGTAAGATGGGTGAATGAGACG 900
Db      3355 GTGGTGGGAGCAGCCGATGTTTCGGGCGCCCTTACGTAAGATGGGTGAATGAGACG 3414
Qy      901 GAGTGTCTGTTCTTCAACACAGCGGCGCCGACAGGGGCAATGCTTCCGCTGTACATGG 960
Db      3415 GAGTGTCTGTTCTTCAACACAGCGGCGCCGACAGGGGCAATGCTTCCGCTGTACATGG 3474
Qy      961 ATGATATGACACCGGGGTTTCAACAGAGCGTGTGGGGGCGCCCGCTGACATCGGGGGTCT 1020
Db      3475 ATGATATGACACCGGGGTTTCAACAGAGCGTGTGGGGGCGCCCGCTGACATCGGGGGTCT 3534
Qy      1021 GGGACACACACTTGTGATCTGCGCCACAGGACTGCTTCGGAAGACATCCCGAGGCACTTAC 1080
Db      3535 GGGACATTAACACTTGTGACTGCGCCACAGGACTGCTTCGGAAGACATCCCGAGGCACTTAC 3594
Qy      1081 ACCAATGCGGTTTCGGGGCTTGGTTG 1107
Db      3595 ACCAATGCGGTTTCGGGGCTTGGTTG 3621

```

## RESULT 5

```

US-09-539-601-25
/ Sequence 25, Application US/09539601C
/ Patent No. 6630343
/ GENERAL INFORMATION:
/ APPLICANT: Bartenhagel, Ralf FW
/ TITLE OF INVENTION: Hepatitis C virus Cell Culture System
/ FILE REFERENCE: all sequences
/ CURRENT APPLICATION NUMBER: US/09/539,601C
/ EARLIER FILING DATE: 2001-08-30
/ EARLIER FILING DATE: 1999-04-03
/ NUMBER OF SEQ ID NOS: 51
/ SOFTWARE: Patent Ver. 2.1
/ SEQ ID NO 25
/ LENGTH: 11076
/ TYPE: DNA
/ ORGANISM: Hepatitis C virus
/ FEATURE:
/ NAME/KEY: 5'UTR
/ LOCATION: (1)..(3411)
/ OTHER INFORMATION: construct I389/Core-3'/5.1
/ NAME/KEY: CDS
/ LOCATION: (342)..(1193)

```

```

/ OTHER INFORMATION: hepatitis C virus core - neomycin
/ OTHER INFORMATION: phosphotransferase fusion protein
/ FEATURE:
/ NAME/KEY: RBS
/ LOCATION: (1202)..(1812)
/ OTHER INFORMATION: internal ribosome entry site from
/ OTHER INFORMATION: encephalomyocarditis virus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1813)..(10845)
/ OTHER INFORMATION: hepatitis C virus polypeptide from core to
/ OTHER INFORMATION: nonstructural protein NS5B; carries cell
/ OTHER INFORMATION: culture-adaptive mutations of clone 5.1
/ FEATURE:
/ NAME/KEY: 3'UTR
/ LOCATION: (10846)..(11076)
/ US-09-539-601-25

Query Match      83.8%; Score 927.8; DB 4; Length 11076;
Best Local Similarity 89.9%; Pred. No. 4,6e-254;
Matches 995; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy      1 TCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCAAGACGCCACATCCCC 60
Db      2515 TCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCAAGACGCCATGCCCTCC 2574
Qy      61 ACTGCGACATATGACAGCCACGTCGATTTGCTGTGGGGGGGCTGCTTCTGCTCCGT 120
Db      2575 ACTGCGACATATGACAGCCACGTCGATTTGCTGTGGGGGGGCTGCTTCTGCTCCGT 2634
Qy      121 ATGATAGTGGGGGATCTCTGCGGATCTGTTTCTGCTGCTCAGCTTTCACTTCTCG 180
Db      2635 ATGATAGTGGGGGATCTCTGCGGATCTGTTTCTGCTGCTCAGCTTTCACTTCTCG 2694
Qy      181 CCTGCGGACATGACAGCGGTACAGACTGCAATGTTCAATCTATCCCGGCACTATCA 240
Db      2695 CCTGCGGACATGACAGCGGTACAGACTGCAATGTTCAATCTATCCCGGCACTATCA 2754
Qy      241 GGTACCGCATGCTGCTGGGATATGATGATGAACTGCTACCTTACACAGCCCTTATGTTGA 300
Db      2755 GGTACCGCATGCTGCTGGGATATGATGATGAACTGCTACCTTACACAGCCCTTATGTTGA 2814
Qy      301 TCGCAGTACTCGGATATCCCAAGGCTGTGTGACATGTTGAGCGGGGCCCACTGAGGA 360
Db      2815 TCGCAGTACTCGGATATCCCAAGGCTGTGTGACATGTTGAGCGGGGCCCACTGAGGA 2874
Qy      361 GTCCTGCGGCGCTTCTGCTATATTCATGTTGGGGAATCGGCTAAGTCTTGTGTTG 420
Db      2875 GTCCTGCGGCGCTTCTGCTATATTCATGTTGGGGAATCGGCTAAGTCTTGTGTTG 2934
Qy      421 ATCTACTCTTTTCCCGCGCTTGAACGGGGAACCTTACACAGAGGGGGACACAGGCGCG 480
Db      2935 ATCTACTCTTTTCCCGCGCTTGAACGGGGAACCTTATGTGACAGGGGGAGATGGCCAA 2994
Qy      481 GCGGCCAAGGGGCTTACATCCCTTTCACACAGTGGGCGGCTCAGAAATCCAGCTTGA 540
Db      2995 AACACCTCGGGGATTCGTCTCTTTTCAACCGGGGTCATCCAGAAATCCAGCTTGA 3054
Qy      541 AACACCAAGGGGAGCTGGGACATCAACAGAACTGCTGAATGCAATGATCTCCCTCA 600
Db      3055 AACACCAAGGGGAGCTGGGACATCAACAGAACTGCTGAATGCAATGATCTCCCTCA 3114
Qy      601 ACTGGTCTCTTCCCGCGCTGTTCTACACGACAGAGTTCAATGCTCCGATCTCAGAG 660
Db      3115 ACTGGTCTCTTCCCGCGCTGTTCTACAGTGAAGTTCATCTGATGATCCAGAG 3174
Qy      661 CGCATGCGCAGCTGCGCGCCCATTTGACAGAACTGATCAGGGGTGGGGTCCCATCTTAT 720
Db      3175 CGCATGCGCAGCTGCGCGCCCATTTGACAGAACTGATCAGGGGTGGGGTCCCATCTTAT 3234
Qy      721 AATGATCCCAAGGCTTGGACAGAGGCTTATGCTGGGACCTACGACCTCAACCGTGT 780
Db      3235 AATGATCCCAAGGCTTGGACAGAGGCTTATGCTGGGACCTACGACCTCAACCGTGT 3294

```

Oy	781	GGTATCGGCGCCGGTTCAGGTGTGTGGCCAGGTACTGTATTCACTCCAAAGCCCTGT	840
Db	3295	GGTATCGTACCCGCGCGCCAGGTGTGTGTGTCCAGTACTGTCTTACCCCAAGCCCTGT	3354
Oy	841	GTGGTGGGGAAGACCGCATCTGTTTGGGCGCCCTTACTGTACAGATGGGGGTGAGATATGACG	900
Db	3355	GTGGTGGGGAAGACCGACCGGTGTGGGCTCTTACTGTACAGTGTGGGGGGAATATGACG	3414
Oy	901	GACGTGCTGCTTCTTCAACAACAACGCGGCGCGCAACGAGGCAACTGTGTTGGCTGTATCATGG	960
Db	3415	GACGTGCTGCTTCTTCAACAACAACGCGGCGCGCAAGCAACTGTGTTGGCTGTATCATGG	3474
Oy	961	ATGATATGACACCGGGTATCCACAGACGTGTGGGGGCCCCCGGTGCAACATGTGGGGGGGTC	1020
Db	3475	ATGATATGACACTGGGGTATCCACAGACGTGTGGGGGCCCCCGGTGTAAACATGTGGGGGGGATC	3534
Oy	1021	GGCAACAACACTTGTGATCTGCCCCCAGACTGTCTTCCGGAAGCATCCCGAAGCCACTTAC	1080
Db	3535	GGCAATAAACCCTTGACTGTGCCCCCAGACTGTCTTCCGGAAGCATCCCGAAGCCACTTAC	3594
Oy	1081	ACCAATGCGGTTGGGGGCTTGGTTG	1107
Db	3595	ACCAATGTGGTTGGGGGCTTGGTTG	3621

```

RESULT 6
US-09-539-601-31
; Sequence 31, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartschlagger, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct 1389/core-3'/19
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; FEATURE:
; NAME/KEY: encephalomyocarditis virus
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polyprotein from core to
; OTHER INFORMATION: nonstructural protein NS5b; carries cell culture
; OTHER INFORMATION: adaptive mutations from clone no. 19
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
; US-09-539-601-31

```

Query Match	83.8%;	Score 927.8;	DB 4;	Length 11076;
Best Local Similarity	89.9%;	Pred. No. 4,6e-254;		
Matches 995; Conservative	0;	Mismatches 112;	Indels 0;	Gaps 0;

QY	1	TTCTCCCGGCTGCTGGGATGAGCGCTCACTCCACCGCTCGCGGCAAGAGACGGCAGCATCCCC	60
Db	2515	TTCTTCGCCGCTGCTGGGATGAGCGCTCACTCCACCGCTCGCGGCAAGAGACGGCTAGCTCCCC	2574
QY	61	ACTCGACAATATACACGACGCACTGCGATTTGCTCTGTTGGGGCGGCTGCCTTCTGCTCCGT	120
Db	2575	ACTATGAGAGATAGACACCGCATGATGATTTGCTGTTGGGGGCGGCTGCCTTCTGCTCCGCT	2634
QY	121	ATGTACGTGGGGGATCTCTGCGGATCTGTTTTCTCGTCTCTACGCTGTTTCACTTTCTCG	180
Db	2635	ATGTACGTGGGAGATCTCTGCGGATCTGTTTTCTCGTCCGCGACGCTGTTCACCTTTCTCG	2694
QY	181	CCTCGCGGACATCGACGCGTACAGACCTGCATAATGTTCAATCTATCCCGGACGTAATCA	240
Db	2695	CCTCGCGGACGAGACATGACAGACCTGCATAATGTTCAATATATCCCGGACGCTACACA	2754
QY	241	GATCACCGCATGAGCTTGGGATATGATGATGAACTGTACCTACACGACGCCCTAGTGTA	300
Db	2755	GATCACCGCATATGAGCTTGGGATATGATGAACTGTACCTACACGACGCCCTAGTGTA	2814
QY	301	TCGACGCTACTCCGGATATCCCAACAGCTGTGTGSAATGTGTGCGGGGCCCCACTGGGGA	360
Db	2815	TCGACGATTACTCCGGATATCCCAACAGCTGTGTGATATGTGTGCGGGGCCCCATTTGGGGA	2874
QY	361	GTCTCTGGCGGGCTTGCCTACTATTTCCATGCTGGGGGAACTGGGGTATAGTCTTGGTTGTG	420
Db	2875	GTCTCTAGCGGGCTTGCCTACTATTTCCATGCTGGGGGAACTGGGGTATAGTCTTATTTGTG	2934
QY	421	ATGCTACTCTTTTGGCGGCGTGTGACGGGGAACTTATACAGACAGGGGGGAGACACAGCGCGC	480
Db	2935	ATGCTACTCTTTTGGCGGCGTGTGACGGGGGAACTATGTGACAGGGGGGAGACATGGCCAAA	2994
QY	481	GCGGCCCCAGGGCTTTCATCTCTTCTTCAACACTGAGGCGCGGCTCAAGAAATTCACGCTTGT	540
Db	2995	AACACCCCTCGGGATTTACGTCCCTCTTTTCAACCCGGGCTATCCAGAAATTCACGCTTGT	3054
QY	541	AACACCAACGGCACTGGCAATGACAGAACTGCTGAACTGCAATGATCTCCCTCCAA	600
Db	3055	AACACCAACGGCACTGGCAATGACAGAACTGCTGAACTGCAATGATCTCCCTCAAC	3114
QY	601	ACTGGGTTCTTTTGGCGGCTGTTTCTAACGCGACAGGTTCAATGTGCTCGGATGCTCAGAG	660
Db	3115	ACTGGGTTCTTTTGGCGGCTGTTTCTAACGCGATGTCAGAGTTCAACTCTGATGCTCCAGAG	3174
QY	661	CGCATGGCGAGCTGCGGCCCATTTGACCAAGTTTGATCAAGGGTGGGGTCCCATCATTTAT	720
Db	3175	CGCATGGCGAGCTGAGGCCCATTTGACCAAGCTTTCCTCAGGGGTGGGGGCCCATCATTTAC	3234
QY	721	AATGATGTCACAGCGCTGTGACCAAGGCTCTATTTGCTGTGCACTACGCACTTCAACCGGT	780
Db	3235	AATGATGTCACACACTCGGACCAAGGCTCTATTTGTTGTGGCACTACGCACTCCCGGCTGGC	3294
QY	781	GGTATCGGCGCGGGTTGCAGATGATGAGGCCAGATGATCTGTTCACTCCAAAGCCCTGTT	840
Db	3295	GGTATCGATCCCGGGGCGGAGATGATGATGCTCAAGTATCTGCTTCACCCCAAGCCCTGTC	3354
QY	841	GTGTGTGGGAGACGACCGATCTGTTGCGCGCCCCCTTACGTAACAGATGGGATGAGAAATGACG	900
Db	3355	GTGTGTGGGAGACGACCGATCTGTTGCGGCTCCCTACGTACAGTTGGGGGAGAAATGACGCG	3414
QY	901	GACGTGCTGCTTCTTCACACACACGCGGCGCCACAGGGGCAACTGTGTTTCGGCTGTATCAG	960
Db	3415	GACGTGCTGCTTCTTTCACACACCGCGGCGCCGACAGGCAACTGTGTTGCTGTATCATAGG	3474
QY	961	ATGAAATGACACGGGTTTACCAAAACGTTGTGGGGGGCCCCCGGTGCACATCGGGGGGGTCT	1020
Db	3475	ATGAAATGACACTGGGTTTACCAAAACGTTGTGGGGGGCCCCCGGTGTATCATGGGGGGATTC	3534
QY	1021	GGCAACAACACTTTGATCTGCCCCACGGAATGCTTCCGGAAGCAATCCGAGGCACTTAC	1080
Db	3535	GGCAATTAACACTTGAATCTGCCCCACGGAATGCTTCCGGAAGCAATCCGAGGCACTTAC	3594
QY	1081	ACCAAAATCGGTTCCGGGCTTGTGTTG 1107	



Db 3595 ACCAAGTGTGCTCGGGGCTTG 3621

RESULT 7  
US-09-014-416-4  
Sequence 4, Application US/09014416  
Patent No. 6153421  
GENERAL INFORMATION:  
APPLICANT: Yanagi, Masayuki  
APPLICANT: Buhk, Jens  
APPLICANT: Emerson, Susanne U.  
APPLICANT: Purcell, Robert H.  
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND  
FILE REFERENCE: 20264276  
CURRENT APPLICATION NUMBER: US/09/014,416  
CURRENT FILING DATE: 1998-01-27  
EARLIER APPLICATION NUMBER: US 60/053,062  
EARLIER FILING DATE: 1997-07-18  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 4  
LENGTH: 9595  
TYPE: DNA  
ORGANISM: Hepatitis C virus  
US-09-014-416-4

Query Match 82.2%; Score 909.8; DB 3; Length 9595;  
Best Local Similarity 89.0%; Pred. No. 5.9e-249;  
Matches 983; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 3 CTCGCCCTGCTGGGTGAGCGCTCACTCCAGCTCGGGGCGGCAAGAGCCGAGCATCCCGAC 62  
DB 1046 CTCGCCCTGCTGGGTGAGCGCTCACTCCAGCTCGGGGCGGCAAGAGTCGCGCTCCCGAC 1105

QY 63 TCGCAATATACGACGCGCAGCTCGATTGCTGTTGGGCGGCTGCTTGTGCTCCGCTAT 122  
DB 1106 TACGACAAATACGACGCGCAGCTCGATTGCTGTTGGGCGGCTGCTTGTGCTCCGCTAT 1165

QY 123 GTAAGTGGGGGATCTCTGCGGATCTGTTTCTGCTCTCACTGTTCACTTCTCGCC 182  
DB 1166 GTAAGTGGGGGATCTCTGCGGATCTGTTTCTGCTCTCACTGTTCACTTCTCGCC 1225

QY 183 TCGCCCAATACAGCGGTACAGAGTGCATTTGTTCAATATATCCCGGCGGCAAGTATCAG 242  
DB 1226 TCGCCCGGATGAGACAGTGCAGTGCAGTGCATCTCAATATATCCCGGCGGCAAGTATCAG 1285

QY 243 TCACCGCATGCTTGGGATATGATGAACTGTCACCTACGACGAGCCCTAGTGTATC 302  
DB 1286 TCACCGCATGCTTGGGATATGATGAACTGTCACCTACGACGAGCCCTAGTGTATC 1345

QY 303 GCAAGTACTCCGGATCCCAAGCTGTGTAAGATGTGCGGGGCGGCGGCACTGGGAGT 362  
DB 1346 GCAAGTACTCCGGATCCCAAGCTGTGTAAGATGTGCGGGGCGGCGGCACTGGGAGT 1405

QY 363 CCTGGCGGCGCTTGGCTACTATTCATGTGGGGAATGGGCTTAAGTCTTGGTGTAT 422  
DB 1406 CCTGGCGGCGCTTGGCTACTATTCATGTGGGGAATGGGCTTAAGTCTTGGTGTAT 1465

QY 423 GCTACTCTTGGCGGCGCTTGGAGCGGGGAACCTTACAGACGAGGGGAGACAGCGCGCG 482  
DB 1466 GCTACTCTTGGCGGCGCTTGGAGCGGGGAACCTTACAGACGAGGGGAGAGCGCGCGCG 1525

QY 483 CGGCCACGGGCTTACATCCCTCTTCAACCTGGGCGGCTCAGAAAATCCAGCTTGTAA 542  
DB 1526 CACCTCGGGGTTACGTCCTTTTCTCATCTGGGGGCTCAGAAAATCCAGCTTGTAA 1585

QY 543 CACCAACGGGAGTGGGACATCAAGAACTGCTTGAATGATGCTTCTCCAAAC 602  
DB 1586 TACCAACGGGAGTGGGACATCAAGAACTGCTTGAATGATGCTTCTCCAAAC 1645

QY 603 TGGGTTCTTGGCGGCGCTGTTTACAGCAGCAAGGTTCAATGCTCCGAGTGTCAAGG 662

Db 1646 TGGGTTCTTGGCGGCGCTGTTTACGACACAGATTCAATGCTCCGGGTGCCGAGCG 1705

QY 663 CATGGCAGCTGCGCGCCCATTTACAGATTTCATGAGGGGTGGGGTCCCATCATTAATA 722  
DB 1706 CATGGCAGCTGCGCGCCCATTTACAGATTTCATGAGGGGTGGGGTCCCATCATTAATA 1765

QY 723 TGAATCCAGCGGTTGAGACAGAGGCGCTATTGCTGAGCACTAGACACTCAACCGTGTG 782  
DB 1766 TGAATCCAGCGGTTGAGACAGAGGCGCTATTGCTGAGCACTAGACACTCAACCGTGTG 1825

QY 783 TATGTCGCCGCGCTTACAGGTGTGAGCCCAAGTATCTGTTCACTCAAGCCCTGTGT 842  
DB 1826 TATGTCGCCGCGCTTACAGGTGTGAGCCCAAGTATCTGTTCACTCAAGCCCTGTGT 1885

QY 843 GGTGGGAGACGAGCTGTTTGGGCGCCCTTACGTACAGATGGGGTGAATATAGACGA 902  
DB 1886 GGTGGGAGACGAGCTGTTTGGGCGCCCTTACGTACAGATGGGGTGAATATAGACGA 1945

QY 903 CGTGTGCTTCTCAACACAGCGGCGGCGCAAGGCGCACTGGTGGCTGTACATGAT 962  
DB 1946 CGTGTGCTTCTCAACACAGCGGCGGCGCAAGGCGCACTGGTGGCTGTACATGAT 2005

QY 963 GAATGACACCGGTTTACCAAGACGTGTGGGGCGCCCGCTGCAACATCGGGGGGTGCG 1022  
DB 2006 GAATGACACCGGTTTACCAAGACGTGTGGGGCGCCCGCTGCAACATCGGGGGGTGCG 2065

QY 1023 CAAACAACCTTGTATGTGCCCCAGAGTGTCTTCCGGAAGCATCCGAGGCCACTTAAC 1082  
DB 2066 TAAACCCACCTTGTATGTGCCCCAGAGTGTCTTCCGGAAGCATCCGAGGCCACTTAAC 2125

QY 1083 CAAATGCGGTTGCGGGGCTTGTGTTG 1107  
DB 2126 AAATGCGGCTGCGGGGCTTGTGTTG 2150

RESULT 8  
US-09-014-416-6  
Sequence 6, Application US/09014416  
Patent No. 6153421  
GENERAL INFORMATION:  
APPLICANT: Yanagi, Masayuki  
APPLICANT: Buhk, Jens  
APPLICANT: Emerson, Susanne U.  
APPLICANT: Purcell, Robert H.  
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND  
FILE REFERENCE: 20264276  
CURRENT APPLICATION NUMBER: US/09/014,416  
CURRENT FILING DATE: 1998-01-27  
EARLIER APPLICATION NUMBER: US 60/053,062  
EARLIER FILING DATE: 1997-07-18  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 6  
LENGTH: 9599  
TYPE: DNA  
ORGANISM: Hepatitis C virus  
US-09-014-416-6

Query Match 82.2%; Score 909.8; DB 3; Length 9599;  
Best Local Similarity 89.0%; Pred. No. 5.9e-249;  
Matches 983; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 3 CTCGCCCTGCTGGGTGAGCGCTCACTCCAGCTCGGGGCGGCAAGAGCGGAGATCCCGAC 62  
DB 1046 CTCGCCCTGCTGGGTGAGCGCTCACTCCAGCTCGGGGCGGCAAGAGTCGCGCTCCCGAC 1105

QY 63 TCGCAATATACGACGCGCAGCTCGATTGCTGTTGGGCGGCTGCTTGTGCTCCGCTAT 122  
DB 1106 TACGACAAATACGACGCGCAGCTCGATTGCTGTTGGGCGGCTGCTTGTGCTCCGCTAT 1165

QY 123 GTAAGTGGGGGATCTCTGCGGATCTGTTTCTGCTCTCACTGTTCACTTCTCGCC 182

||||| 1166 GTACGTGGGGGATCTCGCGGATCTATTTTCTCTGCTCCAGACTGTTCACCTTCTCGCC 1225  
||||| 183 TGGCCGACATACAGCGGTACAGGACTGACATTTGTAATATATCCCGGCGACGATACG 242  
||||| 1226 TCGCCGCGATGAGACAGTGAAGTGAACCTCTCAATATATCCCGCATATATACG 1285  
||||| 243 TCACCCGATAGCTTGGGATATGATGTAAGCTGTCACTACAGACGCCCTAGTGTATC 302  
||||| 1286 TACCCGATAGCTTGGGATATGATGTAAGCTGTCACTACAGACGCCCTAGTGTATC 1345  
||||| 303 GCAGTACTCCGATCCACAGCTGTCTGTGACATGTGTGGCGGGGCCCATCTGGGAGT 362  
||||| 1346 GCAGTGTCTCCGATCCACAGCTGTCTGTGACATGTGTGGCGGGGCCCATCTGGGAGT 1405  
||||| 363 CCGGGGGGCGCTTGCCTACTATTCATGAGGGGAACTGGGCTAAGGCTTGTGTGTAT 422  
||||| 1406 CCGGGGGGCGCTTGCCTACTATTCATGAGGGGAACTGGGCTAAGGCTTGTGTGTAT 1465  
||||| 423 GCTACTCTTTCGCGGCTTACAGGGGAACTTACAGACAGGGGGGACACAGCGCGCGC 482  
||||| 1466 GCTACTCTTTCGCGGCTTACAGGGGAACTTACAGACAGGGGGGAGGGTGGCGCGCAC 1525  
||||| 483 CGCCACGGGCTTACATCCTCTTCAACCTTGGGCGGCTCAGAAAATCCAGCTTGTAA 542  
||||| 1526 CACCTCGGGGTACAGTCCCTTTCTCATCTGGGGGCGCTCAGAAAATCCAGCTTGTAA 1585  
||||| 543 CACCAAGGAGCTGACATCAACAGAACTGCTTGAATGCAATGACCTCCCAAC 602  
||||| 1586 TACCAAGGAGCTGACATCAACAGAACTGCTTGAATGCAATGACCTCCCAAC 1645  
||||| 603 TGGGTTCTTTCGCGGCTTCTACAGGACAGGTTCAATGCGTCCGATGTCTACAGCG 662  
||||| 1646 TGGGTTCTTTCGCGGCTTCTACAGGACAGGTTCAATGCGTCCGATGTCTACAGCG 1705  
||||| 663 CATGGCAGCTGCGCGCCCATTTGACAGTTGATCAGGGGTGGGTTCCATCACTTAA 722  
||||| 1706 CATGGCAGCTGCGCGCCCATTTGATGCTTGCAGGAGGTTGGGCGCCCATCACTTAA 1765  
||||| 723 TGAATCCACAGGCTTGGACAGAGGCGCTTATGCTGCACTAGGACCTCAACCGTGTG 782  
||||| 1766 TGAATCCACAGGCTTGGATCAGAGGCGCTTATGCTGCACTAGGACCTCAACCGTGTG 1825  
||||| 783 TATGTCGCCCGCTTGGAGGTGTGTGGCCAGTGTACTGTTCATCTCAAGCCCTGTGT 842  
||||| 1826 TGTCTACCCGCGCTCGAGGTGTGTGTCCAAGTATGTTTCAACCCCAAGCCCTGTGT 1885  
||||| 843 GGTGGGACGACCGATGTTTGGCGCGCCCTTACGTACAGATGGGGTGAATGAGACGA 902  
||||| 1886 GGTGGGACGACCGATGTTTGGCGCGCGCCCTTACGTATAGCTGGGGGGAATGAGACGA 1945  
||||| 903 CGTGTCTCTTCTCAACAACAACGCGGCGGCAACGCGGGAATGTTTGGGCTGTACATGAT 962  
||||| 1946 CGTGTCTCTTCTCAACAACAACGCGGCGGCAACGCGGGAATGTTTGGGCTGTACATGAT 2005  
||||| 963 GAATAGCAGCGGCTTCAACAAGAGTGTGGGGGCGCGCGTGCACATCGGGGGGCTCG 1022  
||||| 2006 GAATAGCAGCGGCTTCAACAAGAGTGTGGGGGCGCGCGTGCACATCGGGGGGCTCG 2065  
||||| 1023 CAACAACATTTGATCTGCCCCAGAGCTCTTCCGAGACATCCGAGGCCACTTACAC 1082  
||||| 2066 TAAACCGACCTTATCTGCCCCAGAGCTCTTCCGAGACATCCGAGGCCACTTATCAC 2125  
||||| 1083 CAATAGCGGCTTGGGGGCTTGTGTTG 1107  
||||| 2126 AAATGTGCTCGGGGCGCTGTGTTG 2150

RESULT 9  
US-08-150-204E-96  
; Sequence 96, Application US/08150204E  
; Patent No. 6538126  
; GENERAL INFORMATION:

APPLICANT: CHO, Joong Myung  
LEE, Yong Beom  
PARK, Young Woo  
LIM, Kook Jin  
CHOI, Deog Young  
SO, Hong Seob  
KIM, Chun Hyung  
KIM, Sung Taek  
YANG, Jae Young  
TITLE OF INVENTION: HEPATITIS C DIAGNOSTICS AND VACCINES  
NUMBER OF SEQUENCES: 128  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: YANG, Jae Young  
STREET: 386-1, Doryong-dong, Yuseong-gu  
CITY: Daejeon  
STATE: Daejeon  
COUNTRY: Republic of Korea  
ZIP: 305-340  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5inch, 1.44MB storage  
COMPUTER: IBM PC/pentium  
OPERATING SYSTEM: Windows  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/150,204E  
FILING DATE: 20-Apr-1994  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: KR 91-9510  
FILING DATE: 10-JUN-1991  
APPLICATION NUMBER: KR 91-13601  
FILING DATE: 6-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Shahan Islam, Esq.  
REGISTRATION NUMBER: 32,507  
REFERENCE/DOCKET NUMBER: 2695/FLK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 940-8564  
TELEFAX: (212) 940-8776  
INFORMATION FOR SEQ ID NO: 96  
LENGTH: 9472 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
OTHER INFORMATION: KICV-LBC1, Fig. 2  
SEQUENCE DESCRIPTION: SEQ ID NO: 96  
US-08-150-204E-96  
Query Match 80.7%; Score 893.2; DB 4; Length 9472;  
Best Local Similarity 88.0%; Pred. No. 3,2e-244;  
Matches 973; Conservative 0; Mismatches 133; Indels 0; Gaps 0;  
1 TCCTCCGCTGCTGGTACGCTCACTCCACGCTCGCGCCAAAGAGCCAGCATCCG 60  
1045 TCCTCCGCTGCTGGTGGAGCTTACCTCCAGCTCGCGCCAGAAAGAGCATCCG 1104  
61 ACTGCAACATACAGACGCGCATGCTGCTGTGGGGCGGCTCCTTGTGCTCCGCT 120  
1105 ACTACAGACATTCGAGCGCATGCTGCTGTGGGGGTAGCTCTTGTGTTCCGCT 1164  
121 AATGAGTGGGGGATCTGCGGATGCTGTTTCCCTGCTCAGCTTACCTTCCG 180  
1165 AATGAGTGGGGGATCTGCGGATGCTGTTTCCCTGCTCAGCTTACCTTCCG 1224  
181 CCGCGGACATCAGACGCGTACAGACTGCAATTTGTTCAATCTATCCGCGCACGATCA 240  
1225 CCGCGCGGATACAGCGGTACAGACTGCAATTTGTTCAATCTATCCGCGCGGATCA 1284  
241 GGTACCGGATGGCTTGGGATATGATGAACTGTTCACTTACAGACAGCCCTAGTGT 300

Db 1285 GGTCAACCGCATGGCTGGATATGATGATGAAGTGGCTGCTCAACACAGCCCTAGTGGTA 1344  
 QY 301 TCGAGACTACTCCGAGATCCACAAAGCTGTGTGGAGCATGGTGGGGGGCCCACTGGGGGA 360  
 Db 1345 TCGAGACTACTCCGAGATCCACAAAGCTGTGTGGAGCATGGTGGGGGGCCCACTGGGGGA 1404  
 QY 361 GTCTGGCGGGCCCTTGCTCTACTATTCATGATGGGGGAATGGGCTAAGTCTTGGTTGTC 420  
 Db 1405 ATCTGGCGGGCCCTTGCTCTACTATTCATGATGGGGGAATGGGCTAAGTCTTAAATTGGC 1464  
 QY 421 ATCTACTCTTTTCCCGCGCTTGACCGGGAACTTTACACGACAGGGGGGACACACGGCCGC 480  
 Db 1465 ATCTACTCTTTTCCCGCGCTTGACCGGGAACTTTACACGACAGGGGGGACACAGGTCGG 1524  
 QY 481 GCGGCCACAGGGGCTTACCTCTCTTCAACCTGGGCGGCTCAGAAATTCAGCTTGTGA 540  
 Db 1525 GCGGCTAGCTCGCTTAACGCTCTCTTTAGCTGGCGGGCTTACGACCTCCAGCTCATA 1584  
 QY 541 AACACCAAGCGGAGCTGGACATCAACAGAACTGCTTGAATGCAATGATGATCCCTCCAA 600  
 Db 1585 AACACCAAGCGGAGCTGGACATCAACAGAACTGCTTGAATGCAATGATGATCCCTCCAA 1644  
 QY 601 ACTGGGTTCTTCCCGCGCTGTGTCTACACGACAGTTCAATGGCTCGGATGCTCAGAG 660  
 Db 1645 ACTGGGTTCTTCCCGCGCTGTGTCTACAAATACAGGTTCAACGCTCGGATGCTCCGAG 1704  
 QY 661 CGCATGGCGAGCTGGCGGCGCATTTGACCAAGTTCATGATGATGATGATGATGATGAT 720  
 Db 1705 CGCTTGGCGAGCTGGCGGCGCATTTGATGATGATGATGATGATGATGATGATGATGAT 1764  
 QY 721 AATGATCCCAAGCGGCTTGACACAGAGCCCTATTGCTGACATACGACCTCAACGCTGT 780  
 Db 1765 ACTGAGCTCTATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1824  
 QY 781 GGTATGCTGCGCGCTGTGACAGTGTGTGCGGCAAGTGTACTGTTCACTCAACGCTGTT 840  
 Db 1825 GGTATGCTGCGCGCTGTGACAGTGTGTGCGGCAAGTGTACTGTTCACTCAACGCTGTT 1884  
 QY 841 GTGGTGGGGAGCGACCATGCTTCCGGCGGCGCATACGATGGGGTGGAGATGAGACG 900  
 Db 1885 GCGGTGGGGAGCGACCATGCTTCCGGCGGCGCATACGATGGGGTGGAGATGAGACG 1944  
 QY 901 GACGTGCTGCTTCAACAACACGCGGCGCGCACGGGGCACTGTTCCGCTGTACATG 960  
 Db 1945 GACGTGCTGCTTCAACAACACGCGGCGCGCACGGGGCACTGTTCCGCTGTACATG 2004  
 QY 961 ATGAATAGCAACCGGCTTCAACAAGAGTGTGGGGGCGCGCGTGCACATCGGGGGGCTC 1020  
 Db 2005 ATGAATAGCAACCGGCTTCAACAAGAGTGTGGGGGCGCGCGTGTAACTCGGGGGGCTC 2064  
 QY 1021 GGGAAACAACCTTGTATCTGCCCCAGGATGCTTCCGGAAGCATCCCGAGGCACTTAC 1080  
 Db 2065 GGGAAACAACCTTGTATCTGCCCCAGGATGCTTCCGGAAGCATCCCGAGGCACTTAC 2124  
 QY 1081 ACCAAATGCGGCTCGGGGCGCTTGGTT 1106  
 Db 2125 ACCAAATGCGGCTCGGGGCGCTTGGTT 2150  
 RESULT 10  
 US-09-827-688-6  
 ; Sequence 6, Application US/09827688  
 ; Patent No. 6821955  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ORSON, FRANK  
 ; APPLICANT: KINSEY, BERNA  
 ; APPLICANT: BHOAGAL, BALBIR  
 ; TITLE OF INVENTION: MACROAGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION I  
 ; TITLE OF INVENTION: AGENTS  
 ; FILE REFERENCE: F01949051/10004014  
 ; CURRENT APPLICATION NUMBER: US/09/827,688  
 ; CURRENT FILING DATE: 2001-04-06  
 ; PRIOR APPLICATION NUMBER: 60/195,680

; PRIOR FILING DATE: 2000-04-07  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 6  
 ; LENGTH: 9413  
 ; TYPE: DNA  
 ; ORGANISM: HEPATITIS C  
 US-09-827-688-6  
 Query Match 80.5%; Score 891; DB 4; Length 9413;  
 Best Local Similarity 87.8%; Pred. No. 1,38-243;  
 Matches 972; Conservative 0; Mismatches 135; Indels 0; Gaps 0;  
 QY 1 TCCTCCCGCTGCTGGGTAAGGCTCACTCCACAGCTCGCGGACAGAGCCAGCATCCCC 60  
 Db 1032 TTCTCCCGCTGCTGGGTAAGGCTCACTCCACAGCTCGCGGACAGAGCCAGCATCCCC 1091  
 QY 61 ACTGGCAATACAGACGCGCATGATTTGCTGTTGGGGCGGCTGCTTGTGCTCCGT 120  
 Db 1092 ACCACGACATACAGACGCGCATGATTTGCTGTTGGGGCGGCTGCTTGTGCTCCGT 1151  
 QY 121 ATGTAGTGGGGATCTCTGCGGATCTGTTTCTGTCTCTCAAGTGTTCACCTTCTG 180  
 Db 1152 ATGTAGTGGGGATCTCTGCGGATCTGTTTCTGTCTCTCAAGTGTTCACCTTCTCA 1211  
 QY 181 CCTGCCGACATCAGACGCGTACAGACTGCAATTTGTCAATCTATCCCGGACGATTA 240  
 Db 1212 CCTGCCGATTAAGACGCGTACAGACTGCAATTTGTCAATCTATCCCGGACGATTA 1271  
 QY 241 GGTACCGCATGGCTTGGATATGATGATGAATGCTGTACCTTACAGACGCTTATGTTGA 300  
 Db 1272 GGTACCGCATGGCTTGGATATGATGATGAATGCTGTACCTTACAGACGCTTATGTTGA 1331  
 QY 301 TCGACACTACTCCGATATCCCAACAGCTGTGTGGAGATATGGTGGGGGGCCCACTGGGGA 360  
 Db 1332 TCGACACTACTCCGATATCCCAACAGCTGTGTGGAGATATGGTGGGGGGCCCACTGGGGA 1391  
 QY 361 GTCTGGCGGGCCCTTGCTACTATTCATGATGGGGGACGCGGCTAAGGCTTGGTTGTC 420  
 Db 1392 GTCTGGCGGGCCCTTGCTACTATTCATGATGGGGGACGCGGCTAAGGCTTGGTTGTC 1451  
 QY 421 ATGCTACTTTTCCCGCGCTTGAACGGGGAACCTTACACAGACAGGGGGGACACAGCGCGC 480  
 Db 1452 ATGCTACTTTTCCCGCGCTTGAACGGGGAACCTTACACAGACAGGGGGGACACAGCGCGC 1511  
 QY 481 GCCGCCACAGGGCTTACATCCCTTTCAACACTGCGGCGCTCAGAAATCCAGCTTGTGA 540  
 Db 1512 AGCACCCAGAGCTCTGCTGCTGCTCTCAACAGGCGCATCTCAGAAATCCAGCTGTC 1571  
 QY 541 AACACCAAGCGGAGCTGGACATCAACAGAACTGCTTGAATGCAATGATCCCTCCAA 600  
 Db 1572 AACACCAAGCGGAGCTGGACATCAACAGAACTGCTTGAATGCAATGATCCCTCCAA 1631  
 QY 601 ACTGGGTTCTTCCCGCGCTGTGTCTACACGACAGGTTCAATGCTGCCGATGCTCAGAG 660  
 Db 1632 ACTGGGTTCTTCCCGCGCTGTGTCTACACGACAGGTTCAATGCTGCCGATGCTCAGAG 1691  
 QY 661 CGCATGGCGAGCTGCGGCGCATTTGACAGTGTGATCAGGGGTGGGCTCCATCACTTAT 720  
 Db 1692 CGCATGGCGAGCTGCGGCGCATTTGACAGTGTGATCAGGGGTGGGCTCCATCACTTAT 1751  
 QY 721 AATGATCCCAAGCGCTTGAACAGAGGCGCTTATGCTGACATGACGACCTCAACGCTGT 780  
 Db 1752 AATGATCCCAAGCGCTTGAACAGAGGCGCTTATGCTGACATGACGACCTCAACGCTGT 1811  
 QY 781 GGTATGCTGCGCGCTTGAAGTGTGTGCGGCAAGTGTACTGTTTCACTCAAGCCCTGTT 840  
 Db 1812 GGTATGCTGCGCGCTTGAAGTGTGTGCGGCAAGTGTACTGTTTCACTCAAGCCCTGTT 1871  
 QY 841 GTGGTGGGGAGCGACGATGTTTCCGGCGGCGCTTACGTAAGATGGGGTGAATGAGACG 900  
 Db 1872 GTGGTGGGGAGCGACGATGTTTCCGGCGGCGCTTACGTAAGATGGGGTGAATGAGACG 1931

QY 901 GACGTCTGTTCTCAACACAGCGGCGCCACGCGGCACTGTTGGCTGTACATGG 960  
 DB 1932 GACGTCTGTTCTCAACACAGCGGCGCCCTCAAGGCACTGTTGGCTGTACATGG 1991  
 QY 961 ATGAATAGCAGCGGCTTACCAAGAGGTGTGGGGGCCCCCGTGCACATCGGGGGGCTC 1020  
 DB 1992 ATGACAGCAGCTGGGTTCAACAGAGCGTGGGGGGCCCTCCGCAACATCGGGGGGCTC 2051  
 QY 1021 GGCACACACCTTGTATCTGCCCCACGAGCTGCTCCGGAAGCATCCGAGGCACTTAC 1080  
 DB 2052 GGCACACACCTTGTATCTGCCCCACGAGATGCTTCCGGAAGCACCCCGAGGCACTTAC 2111  
 QY 1081 ACCAATGCGGTTGCGGGGCGCTTGTG 1107  
 DB 2112 ACAAGTGTGGCTCGGGGCGCTTGTG 2138

RESULT 11  
 US-08-324-977-11  
 ; Sequence 11, Application US/08324977  
 ; Patent No. 5747339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OKAYAMA, Hiroto  
 ; APPLICANT: FUKU, Isao  
 ; APPLICANT: MORI, Chisato  
 ; APPLICANT: TAKAMIZAWA, Akahisa  
 ; APPLICANT: YOSHIDA, Iwao  
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
 ; NUMBER OF SEQUENCES: 50  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Armstrong, Westernman, Hattori, Mclelland &  
 ; STREET: 1725 K St. N.W. Suite 1000  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/324,977  
 ; FILING DATE: 18-OCT-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 2-167466  
 ; FILING DATE: 25-JUN-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 2-230921  
 ; FILING DATE: 31-AUG-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 2-305605  
 ; FILING DATE: 09-NOV-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/099,706  
 ; FILING DATE: 30-JUL-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/769,996  
 ; FILING DATE: 02-OCT-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/635,451  
 ; FILING DATE: 28-DEC-1990  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Stevens-Smith, Theresa M.  
 ; REGISTRATION NUMBER: 35,281  
 ; REFERENCE/DOCKET NUMBER: 900703D  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 658-2930  
 ; TELEFAX: (202) 887-0357  
 ; TELEX: 440142  
 ; INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 6039 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 1..6039  
 OTHER INFORMATION: /note: "sequence = 333 - 6371 of  
 OTHER INFORMATION: SEQ ID NO: 1"  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..6039  
 US-08-324-977-11

Query Match 80.1%; Score 886.2; DB 1; Length 6039;  
 Best Local Similarity 87.5%; Pred. No. 2.7e-242;  
 Matches 969; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 1 TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCCAAGGACGCCAGCATCCCC 60  
 DB 703 TCCTCCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCCAAGGACGCCAGCATCCCC 762  
 QY 61 ACTGCACAAATAGACGCGCAGCTGATTTGCTGTTGGGGGCGCTGCTTGTGCTCGCT 120  
 DB 763 ACAGAGACATAGACGCGCAGCTGATTTGCTGTTGGGGGCGCTGCTTGTGCTCGCT 822  
 QY 121 ATGTAGTGGGGATCTTGGCGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 DB 823 ATGTAGTGGGGATCTTGGCGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 882  
 QY 181 CCTCGCCGACATCAGCGGATGAGAGCTGCAATGTTCAATCTATCCGGGCAAGTATCA 240  
 DB 883 CTTCCCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 942  
 QY 241 GGTACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
 DB 943 GGTACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1002  
 QY 301 TGGCAGCTACTCGGATGCCACAGCTGTCGTGAGATGATGATGATGATGATGATGATG 360  
 DB 1003 TGGCAGCTACTCGGATGCCACAGCTGTCGTGAGATGATGATGATGATGATGATGATG 1062  
 QY 361 GTCTGGCGGGCTTCTACTACTATTCATGATGATGATGATGATGATGATGATGATG 420  
 DB 1063 GTCTGGCGGGCTTCTACTACTATTCATGATGATGATGATGATGATGATGATGATG 1122  
 QY 421 ATGCTACTTGTGCGCGCTTGAACGGGAACTTACACGACAGGGGGGACACCGGCCG 480  
 DB 1123 ATGCTACTTGTGCGCGCTTGAACGGGAACTTACACGACAGGGGGGACACCGGCCG 1182  
 QY 481 GCCGCCACGCGCTTACATCCCTCTTCAACCTGCGGCTGCAAAAAATCCAGCTTGA 540  
 DB 1183 ACCACCAACAGGCTCTGTCATGTTGCAAGTGGGCGCTCTCAAAAAATCCAGCTTGA 1242  
 QY 541 AACACCAAGGAGCTGGGACATCAACAGAACTGCTTGAATGCAATGACCTCCCA 600  
 DB 1243 AACACCAAGGAGCTGGGACATCAACAGAACTGCTTGAATGCAATGACCTCCCA 1302  
 QY 601 ACTGGGTTCTTCCCGCGCTGTTCTTCAACGACAGGTTCAATGCGTCCGATGCTAG 660  
 DB 1303 ACTGGGTTCTTCCCGCGCTGTTCTTCAACAGGTTCAATGCGTCCGATGCTAG 1362  
 QY 661 CGCATGGCAGCTGCGGCCCATTTGACAGTTGATCAGGGGTGGGGTCCCATCTTAT 720  
 DB 1363 CGCATGGCAGCTGCGGCCCATTTGACAGTTGATCAGGGGTGGGGTCCCATCTTAT 1422  
 QY 721 AATGATCCACGCGCTTGAACAGAGGCGCTATTTGCTGGACCTAGCACTCAACCGTGT 780  
 DB 1423 GCTGATCTAGACATCAGACAGAGGCGCTATTTGCTGGACCTCAACCTCAATGT 1482  
 QY 781 GGTATGTCGCCGCTTGAAGGTGTGTGCGCCAGTGTACTGTTTCACTCAACCGCTGT 840

```

Db 1483 ACCATCGTACCTGGTGGAGGTGTGGCCAGTGTACTGCTTCAACCCCAAGCCCTGTG 1542
QY 841 GTGTGGGAGCAACCCATCTGTTTGGGGCCCTTACGTACAGTGGGGTGAATGAGACG 900
Db 1543 GTGTGGGAGCAACCCATCTGTTTGGGGCCCTTACGTACAGTGGGGTGAATGAGACG 1602
QY 901 GACGTGCTGCTTCAACCAACAGCGGGCCGCAACGGGGCAACTGGTGGCTGTACATGG 960
Db 1603 GACGTGCTGCTTCAACCAACAGCGGGCCGCAACGGGGCAACTGGTGGCTGTACATGG 1662
QY 961 ATGAATAGCAACCGGGTTTACCAAGACGTGTGGGGCCCGCTGTGCAACATCGGGGGGTC 1020
Db 1663 ATGAATAGCAACCGGGTTTACCAAGACGTGTGGGGCCCGCTGTGCAACATCGGGGGGTC 1722
QY 1021 GGCACACACCTTTGATCTGCTCCCAAGACTGCTTCCGGAAGATCCCGAGGCACTTAC 1080
Db 1723 GGCACACACCTTTGATCTGCTCCCAAGACTGCTTCCGGAAGATCCCGAGGCACTTAC 1782
QY 1081 ACCAATGGGGTTGGGGCCCTGGTTG 1107
Db 1783 ACMAATGTGTGTGGGGCCCTGGCTG 1809

RESULT 12
US-08-384-616-11
Sequence 11, Application US/08384616
Patent No. 5847101
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroko
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
NUMBER OF SEQUENCES: 50
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,616
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 6039 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..6039
OTHER INFORMATION: /note: "sequence = 333 - 6371 of
OTHER INFORMATION: SEQ ID NO: 1"
NAME/KEY: CDS
LOCATION: 1..6039
US-08-384-616-11

Query Match 80.1%; Score 886.2; DB 2; Length 6039;
Best Local Similarity 87.5%; Pred. No. 2,7e-242;
Matches 969; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 1 TCCTCCCGCTGCTGGGAGGCTCATCCCAAGCTGCGGCAAGGACCCACATCCCC 60
Db 703 TCCTCCCGCTGCTGGGAGGCTCATCCCAAGCTGCGGCAAGGACCCACATCCCC 762
QY 61 ACTGCACAAATGACAGCCGCAAGTGTGCTGTTGGGGCGGCTGCTTGTGCGCT 120
Db 763 ACCACACGATGACAGCCGCAAGTGTGCTGTTGGGGCGGCTGCTTGTGCGCT 822
QY 121 ATGTAGTGGGGAGATCTGCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 823 ATGTAGTGGGGAGATCTGCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTG 882
QY 181 CTGCGCGACATGACAGCGGATGACAGCTGCAATGTTTCAATCATCTCCGCGACGATCA 240
Db 883 CTGCGCGCGATGACAGCTGCAATGTTTCAATCATCTCCGCGACGATGATG 942
QY 241 GGTACCGCATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 943 GGTACCGCATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1002
QY 301 TCGCAGTACTCCGATCCCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 1003 TCGCAGTACTCCGATCCCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1062
QY 361 GTCCTGCGGGGCTTGCCTACTATTCATGATGATGATGATGATGATGATGATGATG 420
Db 1063 GTCCTGCGGGGCTTGCCTACTATTCATGATGATGATGATGATGATGATGATGATG 1122
QY 421 ATGCTACTTTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 1123 ATGCTACTTTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1182
QY 481 GCCGCCAGCGGCTTACATCTCTTCAACCTGCGGCGGCTGCAAAAATCCAGCTTGA 540
Db 1183 ACCACCAAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1242
QY 541 AACACCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 1243 AACACCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1302
QY 601 ACTGGGTTCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 1303 ACTGGGTTCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362
QY 661 CGCATGGCAGCTGCGGCGGCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 1363 CGCATGGCAGCTGCGGCGGCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1422
QY 721 AATGATGCCAAGGCTTGGACAGAGGCGCTATGCTGGGACTAGCAGCCTCAACGCTGT 780

```

Db 1423 GCTGAGCTTAGCAGATCAGACGAGCCATATTTGCTGGACATACCACTCCACATATG 1482  
Qy 781 GGTATGTGTCGGGGTTGAGGTTGAGGCGCCAGTACCTTCACTCCAAAGCCCTGTT 840  
Db 1483 ACCATGTACTCTGCTGAGGAGTGGGCGCCAGTGTACTGCTTCAACCCAAAGCCCTGTC 1542  
Qy 841 GTGTGTGGGAGCAGACCATCTGTTTCGGGCGCCCTTACGTACAGATGGGAGTGAATGAGACG 900  
Db 1543 GTGTGTGGGAGCAGACCATCTGTTTCGGGAGTGGGCGCCAGTGTACTGCTTCAACCCAAAGCCCTGTC 1602  
Qy 901 GAGTGTGCTCTTCTCAACAACAGCGGCGCCAGCAGGAGCACTGCTTCCGCTTACATG 960  
Db 1603 GAGTGTGCTCTCTCAACAACAGCGGCGCCAGCAGGCAACTGCTTCCGCTTACATG 1662  
Qy 961 ATGATATGACACGGGTTTCAACAAGAGTGTGGGGGCGCCCGTGAACATCGGGGGGGTCT 1020  
Db 1663 ATGATATGACACGGGTTTCAACAAGAGTGTGGGGGCGCCCGTGTACATCGGGGGGGTCT 1722  
Qy 1021 GGCACAACACTTGTGATCTGCCCCCAGCAGACTGCTTCCGAGAGCATCCCGACACTTAC 1080  
Db 1723 GGCACAACACTTGTGATCTGCCCCCAGCAGACTGCTTCCGAGAGCATCCCGACACTTAC 1782  
Qy 1081 ACCAATGCGGTTTGGGGGCTTGGGTTG 1107  
Db 1783 ACAAAATGTGTTTGGGGGCTTGGGCTG 1809

RESULT 13  
US-08-904-686A-11  
Sequence 11, Application US/08904686A  
Patent No. 5998130  
GENERAL INFORMATION:  
APPLICANT: OKAYAMA, Hiroko  
APPLICANT: FUKU, Isao  
APPLICANT: MORI, Chisato  
APPLICANT: TAKAMIZAWA, Akahisa  
APPLICANT: YOSHIDA, Iwao  
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westernman, Hattori, Mclelland &  
ADDRESSEE: Naughton  
STREET: 1725 K St. N.W. Suite 1000  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/904,686A  
FILING DATE: 01-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/324,977  
FILING DATE: 18-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-167466  
FILING DATE: 25-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-230921  
FILING DATE: 31-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-305605  
FILING DATE: 09-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/099,706  
FILING DATE: 30-JUL-1993  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/769,996  
FILING DATE: 02-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/635,451  
FILING DATE: 28-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Mclelland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 900703G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2930  
TELEFAX: (202) 887-0357  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6039 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA from genomic RNA  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..6039  
OTHER INFORMATION: /note: "sequence = 333 - 6371 of  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..6039  
US-08-904-686A-11

Query Match 80.1%; Score 886.2; DB 2; Length 6039;  
Best Local Similarity 87.5%; Pred. No. 2,7e-242;  
Matches 969; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy 1 TCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCCAAGAGCCAGCATCCCC 60  
Db 703 TCCTCCGCTGCTGGGTAGCGCTCACTCCACGCTGCGGCCAAGAGCATCCCC 762  
Qy 61 ACTGCACATAGCAGACGCGCATGATTTGCTGTTGGGCGGCTGCTTCTGCTCGCT 120  
Db 763 ACCAGACATAGCAGACGCGCATGATTTGCTGTTGGGCGGCTGCTTCTGCTCGCT 822  
Qy 121 ATGTACGTGGGGATTTCTGCGGATCTGTTTCTGCTCTCTCACTGTTCACTTCTG 180  
Db 823 ATGTACGTGGGGATTTCTGCGGATCTGTTTCTGCTCTCTCACTGTTCACTTCTG 882  
Qy 181 CCTGCCGACATGACAGCGGACGAGCTGCAATGTTCAATCTATCCGCGCACATGATCA 240  
Db 883 CCTGCCGACATGACAGCGGACGAGCTGCAATGTTCAATCTATCCGCGCACATGATCA 942  
Qy 241 GGTACCGCATGCTTGGATATGATGATGAATGCTGTCACTACAGACCCCTAGTGTGA 300  
Db 943 GGTACCGCATGCTTGGATATGATGATGAATGCTGTCACTACAGACCCCTAGTGTGA 1002  
Qy 301 TGGCAGTACTCCGATCCCAAGCTGTCTGAGCATGTGGCGGGGCCCACTGGGGA 360  
Db 1003 TGGCAGTACTCCGATCCCAAGCTGTCTGAGCATGTGGCGGGGCCCACTGGGGA 1062  
Qy 361 GTCTGCGGGGCTTGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATG 420  
Db 1063 GTCTGCGGGGCTTGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATG 1122  
Qy 421 ATGCTACTTTTGGCGGCTTGAAGGGAACCTTACACGACAGGGGGGACACAGCGCCG 480  
Db 1123 ATGCTACTTTTGGCGGCTTGAAGGGAACCTTACACGACAGGGGGGACACAGCGCCG 1182  
Qy 481 GCCGCCACGGGCTTACATCCTCTTCAACTGCTGGCGCGCTCAAAAATCCAGCTTGA 540  
Db 1183 ACCACCAAGGCTCGTTCATAGTTCGCAAGTGGGCGCTTCAGAAAATCCAGCTTGA 1242  
Qy 541 AACACCAAGGCTCGTTCATAGTTCGCAAGTGGGCGCTTCAGAAAATCCAGCTTGA 600  
Db 1243 AACACCAAGGCTCGTTCATAGTTCGCAAGTGGGCGCTTCAGAAAATCCAGCTTGA 1302

QY 601 ACTGGGTTCTTGGCCGCTGTTCTACACGACAGGTTCAATGCTCGGATGCTCAGAG 660  
 DB 1303 ACTGGGTTCTTGGCCGCTGTTCTACACAGATAGTTTCAACTGCTCGGATGCTCAGAG 1362  
 QY 661 CGCATGGCCAGCTGCGCCGCTTACAGCTTGAATGAGGGGTGGGGTCCCATCTTAT 720  
 DB 1363 CGCATGGCCAGCTGCGCCGCTTACAGCTTGAATGAGGGGTGGGGTCCCATCTTAT 1422  
 QY 721 AATGAGTCCAGGCTTGGACGACGAGGCGCTATTTGCTGGAGCACTACGACCTCAACCGTGT 780  
 DB 1423 GCTGAGTCTAGACGATACGACGAGGCAATTTGCTGGAGCACTACGACCTCAACCAATGT 1482  
 QY 781 GGTATGCTGCGCCGCTTGGACGAGTGTGCGCCGCTTGAATGAGGGTGGGGTCCCATCTTAT 840  
 DB 1483 ACATGCTGCTGCTGCTGCTGAGGAGTGTGCGCCGCTTGAATGAGGGTGGGGTCCCATCTTAT 1542  
 QY 841 GTGCTGGGAGCAGACCGATGCTTGTGCGCCGCTTGAATGAGGGTGGGGTCCCATCTTAT 900  
 DB 1543 GTGCTGGGAGCAGACCGATGCTTGTGCGCCGCTTGAATGAGGGTGGGGTCCCATCTTAT 1602  
 QY 901 GAGTGTGCTTCTCAACAGACGCGCGCGCAACGCGGCGCACTGTTGCGCTGTACATGG 960  
 DB 1603 GAGTGTGCTTCTCAACAGACGCGCGCGCAACGCGGCGCACTGTTGCGCTGTACATGG 1662  
 QY 961 ATGATATGACACCGGCTTCAACAGACGAGTGTGAGGGGCGCCGCTTGAATGAGGGTGGGGTCCCATCTTAT 1020  
 DB 1663 ATGATATGACACCGGCTTCAACAGACGAGTGTGAGGGGCGCCGCTTGAATGAGGGTGGGGTCCCATCTTAT 1722  
 QY 1021 GGCACAAACACTTTGATCTGCGCCGACGAGTGTGCGAAGGATCCGAGGCGCACTTAC 1080  
 DB 1723 GGCACAAACACTTTGATCTGCGCCGACGAGTGTGCGAAGGATCCGAGGCGCACTTAC 1782  
 QY 1081 ACCAATGCGGTTGCGGCGCTTGTGTTG 1107  
 DB 1783 ACCAATGCGGTTGCGGCGCTTGTGTTG 1809

# RESULT 14

US-09-315-850-11  
 Sequence 11, Application US/09315850  
 Patent No. 6217872  
 GENERAL INFORMATION:  
 APPLICANT: OKAYAMA, Hiroto  
 APPLICANT: FUKU, Isao  
 APPLICANT: MORI, Chisato  
 APPLICANT: TAKAMIZAWA, Akahisa  
 APPLICANT: YOSHIDA, Iwao  
 TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
 NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Aramstrong, Westerman, Hattori, Mclelland &  
 ADDRESSEE: Naughton  
 STREET: 1725 K St. N.W. Suite 1000  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 in, 1.44MB  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/315,850  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/904,686  
 FILING DATE: 01-AUG-1997  
 APPLICATION NUMBER: US 08/324,977  
 FILING DATE: 18-OCT-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-167466

FILING DATE: 25-JUN-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-230921  
 FILING DATE: 31-AUG-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-305605  
 FILING DATE: 09-NOV-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/099,706  
 FILING DATE: 30-JUL-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/769,996  
 FILING DATE: 02-OCT-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/635,451  
 FILING DATE: 28-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mclelland, Le-Nhung  
 REGISTRATION NUMBER: 31,541  
 REFERENCE/DOCKET NUMBER: 900703G  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 659-2930  
 TELEFAX: (202) 887-0357  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 6039 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 1..6039 /note: "sequence = 333 - 6371 of  
 OTHER INFORMATION: SEQ ID NO: 1"  
 NAME/KEY: CDS  
 LOCATION: 1..6039  
 US-09-315-850-11

Query Match 80.1%; Score 886.2; DB 3; Length 6039;  
 Best Local Similarity 87.5%; Pred. No. 2.7e-242;  
 Matches 969; Conservative 0; Mismatches 138; Indels 0; Gaps 0;  
 QY 1 TCCTCCCGCTGCTGGGTAGCGCTCATCTCCACGCTTGGCGCCCAAGGACCCAGATCCCC 60  
 DB 703 TCCTCCCGCTGCTGGGTAGCGCTCATCTCCACGCTTGGCGCCCAAGGACCCAGATCCCC 762  
 QY 61 ACTGCGACAATAGCAGCGACGTCGATTTGCTGTTGGGGCGGCTGCTTGTGCTCGGCT 120  
 DB 763 ACCAGCAGATAGAGCGACGTCGATTTGCTGTTGGGGCGGCTGCTTGTGCTCGGCT 822  
 QY 121 ATGTAGTGGGGGATCTGCGGATCTGTTTCTGCTGCTCAGCTGTTCACCTTCTCG 180  
 DB 823 ATGTAGTGGGGGATCTGCGGATCTGTTTCTGCTGCTCAGCTGTTCACCTTCTCG 882  
 QY 181 CCTGCGCAGATCAGCGGTGACGAGCTGCAATTGTTCAATCTATCCGCGCCAGATATCA 240  
 DB 883 CCTGCGCAGATCAGCGGTGACGAGCTGCAATTGTTCAATCTATCCGCGCCAGATATCA 942  
 QY 241 GGTACCGGATGGCTTGGGATATGATGAACTGGTCACTACAGCAGCCGCTAGTGTA 300  
 DB 943 GGTACCGGATGGCTTGGGATATGATGAACTGGTCACTACAGCAGCCGCTAGTGTA 1002  
 QY 301 TCGACGATCTCGGATCCCAAGCTGTGATGACATGATGAGCGGGGCGCCAGCTGGAGA 360  
 DB 1003 TCGACGATCTCGGATCCCAAGCTGTGATGACATGATGAGCGGGGCGCCAGCTGGAGA 1062  
 QY 361 GTCTGGGGGCGCTTGTCTTCAATTCATGATGAGGGAATCGGCTTAAGTCTTGTGTTGG 420  
 DB 1063 GTCTGGGGGCGCTTGTCTTCAATTCATGATGAGGGAATCGGCTTAAGTCTTGTGTTGG 1122  
 QY 421 ATGCTACTTTTGGCGGCGTTGACGGGGAACTTACAGACAGGGGGGACACAGCGGCCCG 480

```

Db      1123 ATGCTACTTTTCTGCGCTTGACGGGATATCCACGTGACAGAGGGGGGGCGCAAGCCAAA 1182
QY      481 GCGGCCACAGGGGTTATACATCCCTCTTCAACACCTGGGGCCGGCTAGAAAATCCAGCTTGA 540
Db      1183 ACCACCAACAGGGCTGCTGCTCAATGTTGCAAGTGGGCGGCTCGAANAATCCAGATTATA 1242
QY      541 AACACCAACAGGGAGCTGGACATCAACAGAACTGCTTGAATGCAATGATGATCCCTCAA 600
Db      1243 AACACCAATGAGGATTGGACATCAACAGAACTGCTTGAATGCAATGATGATGATGATG 1302
QY      601 ACTGGGTTCTTGGCCGCTGTTCTACACGACAGGTTCAATGCTGCTGCTGCTGCTGCTG 660
Db      1303 ACTGGGTTCTTGGCCGCTGTTCTACACAGATGTTCAATGCTGCTGCTGCTGCTGCTG 1362
QY      661 CGCATGGCCAGCTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db      1363 CGCATGGCCAGCTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1422
QY      721 AATGAGTCCACGAGCTTGACACAGAGCCCTATTGCTGCACTACGACCTCAACGCTGT 780
Db      1423 GCTGAGTCTGAGATGACACAGAGCCATATTGCTGCACTACCACTCCACATATG 1482
QY      781 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db      1483 ACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1542
QY      841 GTGTGTGGGAGACGACCATGCTTTCGCGCCCTCAACGACAGATGAGGAGTGAATGAGAG 900
Db      1543 GTGTGTGGGAGACGACCATGCTTTCGCGCCCTCAACGATGAGGAGGAGAAAGAGACT 1602
QY      901 GACGTGCTGCTTCTCAACACAGCGGCGCCGACAGGGGCACTGCTGCTGCTGCTGCTG 960
Db      1603 GACGTGCTGCTTCTCAACACAGCGGCGCCGACAGGGGCACTGCTGCTGCTGCTGCTG 1662
QY      961 ATGAATAGCACCGGGTTACCAAGACGCTGTGTGGGGGCCCCCGTGCACATCGGGGGGCT 1020
Db      1663 ATGAATAGCACCGGGTTACCAAGACATGTGTGGGGGCCCCCGTGTAACTCGGGGGGCT 1722
QY      1021 GGCACCAACCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db      1723 GGCACCAACCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1782
QY      1081 ACCAATGCGGTTGCGGGGCTTGGGTTG 1107
Db      1783 ACAAAATGTGTTGCGGGGCTTGGCTG 1809

```

RESULT 15  
US-08-324-977-13  
Sequence 13, Application US/08324977  
Patent No. 5747339

GENERAL INFORMATION:  
APPLICANT: OKAYAMA, Hiroto  
APPLICANT: FUKU, Isao  
APPLICANT: MORI, Chisato  
APPLICANT: TAKAMIZAWA, Akahisa  
APPLICANT: YOSHIDA, Iwao  
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westerman, Hattori, McIreland &  
ADDRESSEE: Naughton  
STREET: 1725 K St. N.W. Suite 1000  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0

```

?      SOFTWARE: ASCII
?      CURRENT APPLICATION DATA:
?      APPLICATION NUMBER: US/08/324,977
?      FILING DATE: 18-OCT-1994
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: JP 2-167466
?      FILING DATE: 25-JUN-1990
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: JP 2-230921
?      FILING DATE: 31-AUG-1990
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: JP 2-305605
?      FILING DATE: 09-NOV-1990
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: US 08/099,706
?      FILING DATE: 30-JUL-1993
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: US 07/769,996
?      FILING DATE: 02-OCT-1991
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: US 07/635,451
?      FILING DATE: 28-DEC-1990
?      ATTORNEY/AGENT INFORMATION:
?      NAME: Stevens-Smith, Theresa M.
?      REGISTRATION NUMBER: 36,281
?      REFERENCE/DOCKET NUMBER: 900703D
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (202) 659-2930
?      TELEFAX: (202) 887-0357
?      TELEX: 440142
?      INFORMATION FOR SEQ ID NO: 13:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 9030 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: cDNA from genomic RNA
?      FEATURE:
?      NAME/KEY: misc feature
?      LOCATION: 1..9030
?      OTHER INFORMATION: /note: "sequence = 333 - 9362 of
?      OTHER INFORMATION: SEQ ID NO: 1"
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 1..9030
?      US-08-324-977-13

```

Query Match 80.1%; Score 886.2; DB 1; Length 9030;  
Best Local Similarity 87.5%; Pred. No. 3.1e-242;  
Matches 969; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

```

QY      1 TCCCTCCGCTGCTGGGTAGGCTCACTCCGACGCTGGCGGCAAGAGCCGACATCCCC 60
Db      703 TCCCTCCGCTGCTGGGTAGGCTCACTCCGACGCTGGCGGCAAGAGCTCACTCACTCCC 762
QY      61 ACTGCACAAATAGCAGCGCCAGCTGATTTGCTGTGTGGGGCGGCTGCTTGTGCTCGCT 120
Db      763 ACCAGACGATAGCAGCGCCAGCTGATTTGCTGTGTGGGGCGGCTGCTTGTGCTCGCT 822
QY      121 ATGTACGTGGGGGATCTTGGCGGATTTGTTTCTGCTGCTCTCACTGTTCACTTTCTG 180
Db      823 ATGTACGTGGGGGATCTTGGCGGATTTGTTTCTGCTGCTCTCACTGTTCACTTTCTG 882
QY      181 CCTCGCCGACATAGACGCGGACGACGCTGCAATTTGTTCAATCTATCCGGGACGATCA 240
Db      883 CCTCGCCGACATAGACGCGGACGACGCTGCAATTTGTTCAATCTATCCGGGACGATCA 942
QY      241 GGTCAACGCAATGCTTGGATATGATGAACTGTGTCACTTACAGACAGCCCTAGTGTGA 300
Db      943 GGTCAACGCAATGCTTGGATATGATGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1002
QY      301 TGGCAGCTACTCGGATCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360

```



Db 1003 TCCGAGTAATCCGGAATCCCAAGCCGTGTGACATGCTGGCGGGGCCCATGTGGGA 1062  
QY 361 GTCTGGCGGGCCTTCTACTATTCATGATGAGGGAACGGGCTAAGTCTTGTTG 420  
Db 1063 GTCTGGCGGGCCTTCTACTATTCATGATGAGGGAACGGGCTAAGTCTTGTTG 1122  
QY 421 ATGCTACTTCTTGGCCGGCTTGAACGGGGAACCTTACACGACGAGGGGACACAGGCCG 480  
Db 1123 ATGCTACTTCTTGGCCGGCTTGAACGGGGAATCCACGTGACAGGGGGGGCGCAAGCCAA 1182  
QY 481 GCCGCCACGGGCTTACATCCCTCTTCAACCTGGCCGGCTCAGAAAATCAAGCTTGA 540  
Db 1183 ACCACCAACAGGGCTCGTTCATGTTGCAAGTGGGCGCTCAGAAAATCAAGCTTGA 1242  
QY 541 AACACCAACAGGGCTCGTTCATGTTGCAAGTGGGCTCAGAAAATCAAGCTTGA 600  
Db 1243 AACACCAACAGGGCTCGTTCATGTTGCAAGTGGGCTCAGAAAATCAAGCTTGA 1302  
QY 601 ACTGGTCTCTTGGCCGGCTTGAACGGGGAACCTTACACGACGAGGTTCAATGCTCCGATGCTCAG 660  
Db 1303 ACTGGTCTCTTGGCCGGCTTGAACGGGGAACCTTACACGATGTTCAACTGTCGGGTGCCAGAG 1362  
QY 661 CGCATGGCCAGTCCGCCCATTTGACAGCATGATCAGGGGTGGGGTCCCATCACTAT 720  
Db 1363 CGCATGGCCAGTCCGCCCATTTGACAGCATGATCAGGGGTGGGGTCCCATCACTAT 1422  
QY 721 AATGATCCCAAGGCTTGAACGAGGCTTATGCTGACATAGCACTCAACCTGT 780  
Db 1423 GCTGATCTAGCAGATCAGACGAGGCTATTTGCTGCACTACCCACTCAATAT 1482  
QY 781 GGTATGTCGCCCGTTCAGAGTGTGGGCCAGTGTACTGTTCACTCCAGCCCTGT 840  
Db 1483 ACCATGTAATCTGTCGAGAGTGTGGGCCAGTGTACTGTTCACTCCAGCCCTGT 1542  
QY 841 GTGTTGGGGAAGACCGATCGTTTCGGCGCCCTTACGTACAGATGGGGTGAATGAGACG 900  
Db 1543 GTGTTGGGGAAGACCGATCGTTTCGGCGCCCTTACGTACAGATGGGGTGAATGAGACG 1602  
QY 901 GACGTGCTGCTTCAACACAGCGCGGCCCACTGTTCCGCTGTACATG 960  
Db 1603 GACGTGCTGCTTCAACACAGCGCGGCCCACTGTTCCGCTGTACATG 1662  
QY 961 ATGAATAGCACCGGTTCAACAGCGTGTGGGGGCCCGCCGTCACATCGGGGGGTC 1020  
Db 1663 ATGAATAGCACCGGTTCAACAGCATGTGGGGGCCCGCCGTCATCGGGGGGTC 1722  
QY 1021 GGCACACACATTTGATCTGCCCAAGACTGCTTCGGAAGCATCCCGAGCCACTTAC 1080  
Db 1723 GGCACACACATTTGATCTGCCCAAGACTGCTTCGGAAGCATCCCGAGCCACTTAC 1782  
QY 1081 ACCAAATGCGGTTCCGGGCTTGGTTG 1107  
Db 1783 ACAAATGCGGTTCCGGGCTTGGCTG 1809

Search completed: February 21, 2005, 09:05:30  
Job time : 231 secs

**THIS PAGE BLANK (USPTO)**



OY	154	TTGCTCTCTGACGTTTTCACCTTTCGCTCGCCGACATGACAGGATACAGGACTGCAAT	213
Db	412	CTTGAGATTACGCTGATATCA---TCTGGCTTACAGACCATGATGTTTGCAATTAATGCAAC	356
OY	214	TGTTCAATCTATTCGCCGACGACGTAATCAGSTCACCGCATG--GCTTGGATATGATGATGAA	272
Db	355	TGCTTCATCTATTCCTGGGGCCATCACTGSAACCGTAATGAGCATATGGGACATGATGATGAA	296
OY	273	CTGCTACCTTACAGCAGCCCTAGTGGTATGCAAGCTACTCCGGATTCCTCAACAAGCTGTGGT	332
Db	295	CTGGTCGTCGACCCGCTGCTATGATCATATGCGCTAGCGCAATGCGCCTTCTTGAGCTCATCAT	236
OY	333	GGAATAGTGTGGCGGGGGGCCCATCTGGGAGATCTCGCGGGCCCTTGCCTATATTCATAGST	392
Db	235	AGATTATCTACAGCCGGGGCTCACTGGGGGGGTATGTTGCGGCTTAACTCATCTTCTATGCA	176
OY	393	GGGGAACCTGGGCTAAGATCTTGATGTGATGCTACTCTTTCACCGGCTTGACGGGGAAAC	452
Db	175	GGGAGCGTGGCGGAAAGTCGCTGTCTATCTCTTGTAAACCGCTGGGGTACACGGGAGCAC	116
OY	453	TTTACACGACGAGGGGGAC-ACACGGCGCGGCTGCCACGGGCTTATCATCTCCTTTTCACAC	511
Db	115	CCAGGTCACAGATGGCACTGCTAAGGCGCGCCGACAGAGTTTGTGTCAGCTTGCTCAACC	56
OY	512	CTGGGCGGGCTCAAGAAATTCAGCTTTTAAACCAACAGGACGCTGG	558
Db	55	CTGGGCTTAAACAGAAACATCCAATCATATTAGATCTNATTTGCAATGGG	9

RESULT 2	AV758366/c	492 bp	mRNA	linear	EST 19-OCT-2000
LOCUS	AV758366				
DEFINITION	AV758366 BM Homo sapiens cDNA clone BMEKA03 5', mRNA sequence.				

VERSION	AV758366.1	GI:10916214
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE  
AUTHORS  
1 (bases 1 to 492)  
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,

Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.

TITLE	Homo sapiens cDNA BM clones
JOURNAL	unpublished (2000)
COMMENT	Contact: Zeguang Han

201203, P. R. China  
Tel: 86-21-50801919 (ex.45)  
Fax: 86-21-50801000

Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES	Location/Qualifiers
source	1. .492

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone="BMPAKA03"
/tissue_type="bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM2.8"
/clone_id="BM"
/note="Vector: pTriplEx2; Site_1: sf1A; Site_2: sf1B"

```

Query Match	9.4%	Score 104.4	DB 1	Length 492
Best Local Similarity	56.1%	Pred. No. 1.2e-17		
Matches 254; Conservative	0	Mismatches 194	Indels 5	Gaps 3

Oy	110	TCAGCTCCGCTAATGATAGTGGGGGATCTTCGCGGATCTGTTTCTCTGCTCCAGCTGT	169
Db	457	TGTATCAGCTCACTACGTGTGGACCTCTGCGTGGGGGTATCGCTTGACGCCCACTGA	398
Oy	170	TCACCTCTTCGCGCTGCGCGACATCAGACGGGTACAGACTGCAATTGTTCAATCTATCCG	229
Db	397	TTA---TCTCTCAGACCAACAATGTTGTTGTGCAAGATCAACTGCTCATTTCTATCTCG	341
Oy	230	GCCACGATATCAGCTC-AACCGATGCGCTTGGGATATGATGATGAACTGGTCACTCAACGA	288
Db	340	GCTCAATCAGCTGAGCTACAGTATGGCAATGAGCTATGATGATGAACTGTGGCAGCCGCT	281
Oy	289	GCCCTAGTGTATGCAAGCTACTCCGATCCCAACAGCTGTGTGACATGTGTGCGGG	348
Db	280	TTCATGATACGTGGGGTACGGCAATGCGGCTTCTGAACTGCTCATAGATATCATTAAGCTGG	221
Oy	349	GCCCACTGGGAGTCTCTGGCGGGGCTTGCTCTATTCATTCATGTGTGGGAACTGGGCTAAG	408
Db	220	GCACACTCGCGCGGTCAATGTTGCGGTACGTTACTTCACAAATGACGGGAGCGTTGGCCAAA	161
Oy	409	GTTCTGGTGTGATGATCTACTCTTGTTCGCGGCGTTGACCGGGAACCTTACACGACAGGGG	468
Db	160	GTACTGTCACTCATCTGTTCATCCGCTGTGGGGTTATGACGAAACCAAGTATACAGTGGT	101
Oy	469	ACACAGCGCGCGCGGCCCAAGGGCT-TACATCCCTTTTCAACACTTGGGCGCGGCTCAGAA	527
Db	100	ATCACTCGCGCGCGACGACAGCATTTGTGTGATGATATGCTGCGCTCAGACAGTA	41
Oy	528	AATCAGCTTGAACAACAACGCGCAGCTGGCA	560
Db	40	CATCCAGCTCATGANNCGTAAATGGCAATGAGNA	8

RESULT 3			
R28798	179 bp	mRNA	linear
LOCUS			EST 03-DEC-1999
DEFINITION	F0-266D 22 week old human fetal liver	cdna library	Homo sapiens
	cdna clone F0-266D 5'	mRNA sequence.	

KEYWORDS	EST.
SOURCE	Homo sapiens (human)

REFERENCE  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 179)  
Chel E; Cho Y C; Chung Y C and Chiu H C

**AUTHORS** Choi, S.S., Yun, J.W., Choi, E.K., Cho, Y.G., Sung, Y.C. and Shun, H.S.  
**TITLE** Construction of a gene expression profile of a human fetal liver by single-pass cDNA sequencing (1995)

PUBMED 8535075  
COMMENT Contact: Hee-Sup Shin  
hansj@ummc.org Genetics

Pohang Institute of Science & Technology  
San31, Hyogadong Pohang, 790-784 Republic of Korea  
Tel. 150-70-0001

**Tel: 562-279-2291**  
**Fax: 562-279-2199**

```

FEATURES
    source
        location/Qualifiers
            1..179
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="FO 26CD"
                /lab_host="XLA-blue MRF"

```

**ORIGIN**

XhoI: The cDNA library made by oligo-dT primed and directionally cloned between 5' XhoI 'XhoI3' sites."

Query Match	5.2%	Score 58	DB 7	Length 179
Best Local Similarity	86.5%	Pred. No. 7e-05		
Match 64	Conservative 0	Mismatches 10	Indels 0	Gaps 0
Db	521	CTCAGAAATTCAGCTTGTAAACACCAACGACAGCTGGACATCAACGAATGCTTGA	580	
Qy	63	CGGAAACATTCAGCTGCTAAACGCTATATGACAGCTGCACATCAACGAGACTGCTTGA	122	
Db	581	ACTGCATGACTCC	594	
Qy	123	ATTGCAATGACTCC	136	
RESULT 4				
CNS004N/c				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
ACCESSION				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
ORIGIN				

Df		716	CSSTTCCSBHSBKSSSGSCSBSCTSSBTGBSTGYTTTTCSSSGCSTCSGGTTCBBSTT	657
Oy		809	GCCCCAGTGATCGTTTCACTTCCAAAGCCCTGTGTGTGGGACGACCGATCGTTCCGCC	868
Df		656	SSSTSSTTSTTSBTVCCTTKCKKCYCTCSYTBTTTTSBDBCSGCSGABTBSCSS597	
Oy		869	CCCTACGTACAGATGGGGTGAGAAATGACACGACGTCGCTTCTCAACAACAGCGGC	928
Df		596	SGSTTSSSSSSSSSKCSTSSSSSTSSSTSSGSGTSSCSTBCCTCCGCSCTSCCCSCCS	537
Oy		929	CGCACGAGGGAACGTGTTCCGCTGTCATCATGATGAATAGAACCGGGTTCCACCAAGCT	988
Df		536	CTCCSCSYSCSTCCSCCTSSSCTCGSCSTSSSCTCTCCTGTGTSBTTBSSSTGSCSS	477
Oy		989	GTCGGGGGCC 998	
Df		476	TSTGTTTTSS 467	
RESULT 5				
CNS0091P				
LOCUS				
DEFINITION		CNS0091P	925 bp DNA linear GSS 03-JUN-1999	
		Drosophila melanogaster genome survey sequence TETJ end of BAC #		
		BACR19D16 of RPECI-98 library from Drosophila melanogaster (fruit		
		fly), genomic survey sequence.		
ACCESSION		AL053013	GI:4934461	
VERSION		AL053013.1		
KEYWORDS		GSS.		
SOURCE		Drosophila melanogaster (fruit fly)		
ORGANISM		Drosophila melanogaster		
		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
		Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
		Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE		1 (bases 1 to 925)		
AUTHORS		Genoscope.		
TITLE		Direct Submission		
JOURNAL		Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;		
		BP 191 91006 EVRY cedex - FRANCE (E-mail : seget@genoscope.cns.fr		
		- Web : www.genoscope.cns.fr)		
COMMENT		Determination of this BAC-end sequence was carried out as part of a		
		collaboration with the Berkeley Drosophila Genome Project (BDGP).		
		The BDGP is constructing a physical map of the Drosophila		
		melanogaster genome using these BACs. For further information		
		please see http://www.fruitfly.org The BDGP Drosophila		
		Melanogaster BAC library was prepared by Kazutoyo Osogawa and		
		Aaron Mammoser in Pieter de Jong's laboratory in the Department of		
		Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,		
		NY. The library is named RPECI-98 and was constructed by partial		
		EcoRI digestion of Drosophila DNA provided by the BDGP from the		
		isogenic strain Y2; cn bw sp, the same strain used for the BDGP's		
		P1 and EST libraries. A more detailed description of the library		
		and how to order individual BAC clones, the entire library, or		
		filters for hybridization from the BACPAC Resource Center can be		
		found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.		
FEATURES				
source		Location/Qualifiers		
		1..925		
		/organism="Drosophila melanogaster"		
		/mol_type="genomic DNA"		
		/db_xref="taxon:7227"		
		/clone="BACR19D16"		
		/clone_1bp="RPECI-98"		
		/note="end : TETj"		
ORIGIN				
Query Match		3.8%; Score 42.4; DB 9; Length 925;		
Best Local Similarity		13.2%; Pred. No.2.2; Indels 0; Gaps 0;		
Matches		44; Conservative 149; Mismatches 141;		
Oy		226 CCGCGCACGATATAGTCAACCGATGGCTTGGGATATGATGAACGTGTCACTTACA	285	
Df		580 CSSSSGCBCCCSCSSGCSBSSBKSSSTSSBSCSCCSSKSVCGTSCSSSSCSS	639	
Oy		286 GCAGCCTTAGTGATGCGACTACTCCGATCCACCAAGCTGTGCGATGATGTGCG	345	

```

Db      640 SSSSTSSSTSSSTSSSSSSSSSSSYTTSKTSASGSGWMSAGGSGSTGSSSS 699
      346 GGGGCCCACTGGGAGCTCCGCGGCGCTTGCCTACTATTCANGTGGGAGTGGGCT 405
      700 SSSSTSSSTSSSSGSSSTSSBSSSSSSSSSSSSSTSSBSCSTSSSSSSSSSTCS 759
Qy      406 AAGCTCTTGTTGATGATCTTCTTTCGCGGCTTGACGGGAACTTACACAGAGG 465
      760 CTCCTCCSYSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTCCSCCCYMTCCSTYB 819
Db      466 GGGACACAGCGCGCGCGCGCCACGCGGCTTACATCCCTTTCACACCTGGCGGCTCAG 525
      820 MBCYTSSTCGSSSSSGKGGVTKCGCGCGSSSTNGMBGTSASACSSSSSSSSSVSSSS 879
Qy      526 AAAATCCAGCTGTAAACACCAACGAGAGCTGGC 559
      880 KSSASSSVSSSGSGVSNSSASAKSSSSGVS 913
Db

RESULT 6
CB651667/c 731 bp  mRNA  linear  EST 08-APR-2003
LOCUS      OSJNB16L12.r OSJNB Oryza sativa (japonica cultivar-group) cDNA
DEFINITION CB651667
VERSION     CB651667
ACCESSION  CB651667.1 GI:29646660
KEYWORDS   EST.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
            1 (bases 1 to 731)
            Jantsauriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
            Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
            Unpublished (2003)
            Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR Primers
            FORWARD: gta aac cga cgg cca gtc
            BACKWARD: gga aac agc tat gac cat g
            Plate: 16 row: L Column: 12
            Seq primer: gga aac agc tat gac cat g.
            Location/Qualifiers
            1..731
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /culivar="Nipponbare"
            /db_xref="taxon:39947"
            /clone="OSJNB16L12"
            /rlnsue_type="leaf"
            /dev_stage="3 week"
            /lab_host="DH10B"
            /clone_lib="OSJNB"
            /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
            XhoI; 24 hrs after inoculation with Rice Blast (Ch
            86061)"

ORIGIN
Query Match      3.8%; Score 42.2; DB 6; Length 731;
Best Local Similarity 55.9%; Pred. No. 2.4;
Matches 80; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy      943 TGGTTCGGCTGATCATGATGATATGACCGGGTTTCAACAGAGTGTGGGGGCCCCCG 1002

```

```

Db      564 TCGTTCGTCGCTTCGCGGACGACGACGCAACCCGAGCCGAAATGAGAGGTG 505
Qy      1003 TGCACATCGGGGGGGTGGCAACAACATTGATCTGCCCCGAGACTGTTCCGGAAG 1062
      504 TGCACATCGAGAGAGTGCATCTTCAACATCTGGGAGACTCTCCATGAGTCTTCTCGCG 445
Qy      1063 CATCCGAGGCCACTTACACCAA 1085
      444 CCAGGACCTCAGAGATGCAA 422
Db

RESULT 7
CNS01219 864 bp  DNA  linear  GSS 01-SEP-2000
LOCUS      Tetraodon nigroviridis genome survey sequence pUC-ori end of clone
DEFINITION 219D07 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION  AL173718
VERSION     AL173718.1 GI:7811775
KEYWORDS   GSS; genome survey sequence.
SOURCE     Tetraodon nigroviridis
ORGANISM   Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodontidae; Tetraodon.
            1
            Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Broctier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
            Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
            Nat. Genet. 25 (2), 235-238 (2000)
            2
            Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
            Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
            Genome Res. 10 (7), 939-949 (2000)
            3 (bases 1 to 864)
            Genoscope.
            Direct Submission
            Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
            Location/Qualifiers
            1..864
            /organism="Tetraodon nigroviridis"
            /mol_type="genomic DNA"
            /db_xref="taxon:99883"
            /clone="219D07"
            /clone_lib="G"
            /note="Genoscope sequence ID : C0AG219CB04SP1-end :
            pUC-ori"

ORIGIN
Query Match      3.7%; Score 41.4; DB 9; Length 864;
Best Local Similarity 42.3%; Pred. No. 4.1;
Matches 150; Conservative 12; Mismatches 193; Indels 0; Gaps 0;

Qy      406 AAGCTCTTGTTGATGATCTTCTTTCGCGGCTTGACGGGAACTTACACAGAGG 465
      88 ACCTGCTGACCTGAGAGATATCTCATCTGACACAGAGAGCGCCCTTACGAGTGCCT 147

```

ORIGIN	
Query Match	3.7% Score 40.8; DB 8; Length 628
Best Local Similarity	53.0% Pred. No. 5.6;

**KEYWORDS** GSS,  
**SOURCE** Zea mays  
**ORGANISM** Zea mays  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE  
AUTHORS

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 774)

WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennetzen,J.

TITLE  
JOURNAL  
COMMENT

Maize Genomics Consortium  
Unpublished (2003)  
Other GSSs: PUFUQ78TD  
Contact: Cathy WhiteLaw  
TIGR

5712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteLaw@tigr.org

Seq primer: TR  
Class: sheared ends.

## FEATURES

source  
Location/Qualifiers

1..774  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_1lb="ZM\_0.6-1.0\_KB"  
/note="Vector: PCR4-TDPO; Site 1: EcoRI; 0.6-1.0 kb high  
Cot selected genomic DNA library"

## ORIGIN

Query Match 3.7%; Score 40.8; DB 8; Length 774;  
Best Local Similarity 53.0%; Pred. No. 5.9;  
Matches 87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 331 GTGGAATGATGGCGGGGGCCCACTGGGAGATCCTGGCGGCGCTTGTCTTTCATG 390  
DB 425 GTACCCGTGTCCCTGCACTCCACGGGCGAGCCCGCTGCTTGCGAGAGCCCG 484  
QY 391 GTGGGAACTGGGCTAAGTCTTGTGTGATGCTACTCTTTGCCGCGCTTGAACGGG 450  
DB 485 GTGGCGGAGTGTCTACACGACGACATGACCGCGCTGCGCGCGCGCGTGAAGCC 544  
QY 451 CCTTACACGACAGGGGGACACACGCGCGCGCGCGCGCGCT 494  
DB 545 GATGCTGCCGCCGCCGCCGCGCGCGCGCGCGCGCGCGCGCT 588

## RESULT 11

## LOCUS

BZ638670 837 bp DNA linear GSS 29-JAN-2003  
OGCCW777C ZM 0.7-1.5\_KB Zea mays genomic clone ZMMBMA0143N09,  
genomic survey sequence.

## DEFINITION

BZ638670

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 837)  
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Contact: Cathy WhiteLaw  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteLaw@tigr.org  
Seq primer: TP  
Class: sheared ends.

FEATURES  
source

Location/Qualifiers  
1..837  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_1lb="ZMMBMA0143N09"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Query Match 3.7%; Score 40.8; DB 8; Length 837;  
Best Local Similarity 53.0%; Pred. No. 6;  
Matches 87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 331 GTGGAATGATGGCGGGGCCCACTGGGAGATCCTGGCGGCGCTTGTCTTTCATG 390  
DB 14 GTACCCGTGTCCCTGCACTCCACGGGCGAGCCCGCTGCTTGCGAGAGCCCG 73  
QY 391 GTGGGAACTGGGCTAAGTCTTGTGTGATGCTACTCTTTGCCGCGCTTGAACGGG 450  
DB 74 GTGGCGGAGTGTCTACACGACGACATGACCGCGCTGCGCGCGCGCGTGAAGCC 133  
QY 451 CCTTACACGACAGGGGGACACACGCGCGCGCGCGCGCGCT 494  
DB 134 GATGCTGCCGCCGCCGCCGCGCGCGCGCGCGCGCGCGCGCT 177

## RESULT 12

## LOCUS

CC692345 855 bp DNA linear GSS 19-JUN-2003  
OGMBB44TH ZM 0.7-1.5\_KB Zea mays genomic clone ZMMBMA0332H16,  
genomic survey sequence.

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Unpublished (2002)  
Other\_GSSs: OGMBB44TV  
Contact: Cathy WhiteLaw  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteLaw@tigr.org

Seq primer: TR  
Class: sheared ends.

## FEATURES

## source

Location/Qualifiers  
1..855  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_1lb="ZMMBMA0332H16"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Query Match 3.7%; Score 40.8; DB 9; Length 855;  
Best Local Similarity 53.0%; Pred. No. 6.1;  
Matches 87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;



Qy	331 GTGGACATGATGATGCGGGGCGCCACTGCGGGGAAGTCTCGGGCGGGCTTGCTACTATTTCATG	390
Db	852 GTACCCGTTGTTCTCTTGCACTCCACGGGACGGCCCGCTGCTTGCGGAAGAACCCGG	793
Qy	391 GTGGGAACTGGGCTTAAGTCTTGTGTTGATGATGATCTTTTGCCGGCGTTGACGGGAA	450
Db	792 GTGGCGGACGTCGTCAACGACACGAGATGACACCGCTCGCGCGCGGTAGCCTGGA	733
Qy	451 CCTTACACGACAGGGGGGACACACGCGCGCGCCACGGGCT	494
Db	732 GATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGAGGCT	689
RESULT 13		
LOCUS	CC692352	
DEFINITION	CG692352	867 bp DNA linear GSS 19-JUN-2003
ACCESSION	OGMBB44TV ZM 0.7.1.5_KB Zea mays genomic clone ZMMBMA0332H16,	
VERSION	CG692352	genomic survey sequence.
KEYWORDS	CG692352.1 GI:32097128	
SOURCE	GSS.	
ORGANISM	Zea mays	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.	
AUTHORS	1 (bases 1 to 867) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Budiman,V.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.	
TITLE	Unpublished (2002)	
JOURNAL	Other GSSs: OGMBB44TH	
COMMENT	Contact: Cathy Whitelaw	
	9712 Medical Center Drive, Rockville, MD 20850, USA	
	Tel: 301-838-5843	
	Fax: 301-838-0208	
	Email: whitelaw@ligr.org	
	Seq primer: TP	
	Class: sheared ends.	
FEATURES	location/Qualifiers	
source	1..867	
	/organism="Zea mays"	
	/mol_type="Genomic DNA"	
	/strain="B73"	
	/db_xref="taxon:4577"	
	/clone="ZMMBMA0332H16"	
	/clone_1lb="ZM_0.7.1.5_KB"	
	/note="Vector: pBCSK-; Site 1: HincII, 0.7-1.5 kb methylation filtered genomic DNA library"	
ORIGIN		
Query Match	3.7%; Score 40.8; DB 9; Length 867;	
Best Local Similarity	53.0%; Pred. No. 6.1;	
Matches	87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;	
Qy	331 GTGACATGATGATGCGGGGCGCCACTGCGGGGAAGTCTCGGGCGGGCTTGCTACTATTTCATG	390
Db	16 GTACCCGTTGTTCTCTTGCACTCCACGGGACGGCCCGCTGCTTGCGGAAGAACCCGG	75
Qy	391 GTGGGAACTGGGCTTAAGTCTTGTGTTGATGATGATCTTTTGCCGGCGTTGACGGGAA	450
Db	76 GTGGCGGACGTCGTCAACGACACGAGATGACACCGCTCGCGCGCGCGCTGAGCTGGA	135
Qy	451 CCTTACACGACAGGGGGGACACACGCGCGCGCCACGGGCT	494
Db	136 GATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGAGAGGCT	179
RESULT 14		
LOCUS	CC977474	
	897 bp DNA linear GSS 18-AUG-2003	

DEFINITION	ZUAB789TH.ZM.3.0.4.0_KB_Zea_mays_genomic_clone_ZMMBP0012P10, genomic survey sequence.
ACCESSION	CG377474
VERSION	CC377474.1 GI:33837352
KEYWORDS	GSS.
SOURCE	
ORGANISM	Zea mays
REFERENCE	Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 897)
AUTHORS	Whiteleaw,C.A., Quackenbush,J., Van Aken,S., Uteback,T., Renick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomes Unpublished (2002) Other GSSs: ZUAB789TV Contact: Cathy Whiteleaw TIGR
TITLE	
JOURNAL	
COMMENT	9712 Medical Center Drive, Rockville, MD 20850, USA Tel.: 301-838-5843 Fax: 301-838-0208 Email: whiteleaw@tigr.org Seq primer: TR Class: sheared ends.
FEATURES	
Source	Location/Qualifiers 1..897 /organism="Zea mays" /mol_type="genomic DNA" /strain="873" /db_xref="taxon:4577" /clone="ZMMBP0012P10" /clone_1lb="ZM_3.0.4.0_KB" /note="Vector: pBCSK-/ Site_1: HincII; 3-4 kb 'unfiltered' genomic DNA library"
ORIGIN	
Query Match	3.7%; Score 40.8; DB 9; Length 897;
Best Local Similarity	53.0%; Pred. No. 6.1;
Matches	87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
Db	331 GTGACATGTGGGGGGGCCCACTGGAGACTCTGGCGGCCTTGCTACTATTTCATG 390 658 GTACCCTTTCTCCCTGACTCCACGGGACGGCCCGCTGCTGGCGAAGAAGCCCGG 717 391 GTGGGGAACCTGGGCTAAGCTCTTGATGATGCTACTCTTTGGCGGCGCGTTGACGGGAA 450 718 GTGGCGGACGTCGTCAAGCACACGATCGACCGCCTCGCGCGCGCGCGTAGCCCTGGA 777 451 CTTTACACGACAGGGGGGACACACGCGCGCGCGCCCGCACGGGCT 494 778 GATGCTGCG 821
RESULT 15	
LOCUS	CG286439
DEFINITION	914 bp DNA linear GSS 25-AUG-2003
ACCESSION	OC3BE55TH.ZM.0.7.1.5_KB_Zea_mays_genomic_clone_ZMMBHa0767J13,
VERSION	CG286439
KEYWORDS	genomic survey sequence.
SOURCE	CG286439.1 GI:34200653
ORGANISM	GSS. Zea mays Zea mays Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 914) Whiteleaw,C.A., Quackenbush,J., Van Aken,S., Uteback,T., Renick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomes Unpublished (2002)
TITLE	
JOURNAL	
COMMENT	

## COMMENT

Other\_GSSs: OG3BE5TV  
Contact: Cathy WhiteLaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteLaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

## FEATURES

location/Qualifiers  
source 1..914  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0767J13"  
/note="Vector: pBSK-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Query Match 3.7%; Score 40.8; DB 9; Length 914;  
Best Local Similarity 53.0%; Pred. No. 6.2; Mismatches 77; Indels 0; Gaps 0;  
Matches 87; Conservative 0;

QY	331	GTGACATGCTGGCGGGGCCCACTGGGAGTCTTGCGGCTTGTCTACTATTCCATG	390
DB	116	GTACCGGTGTCTCTGACTCCACGGGCAAGCCCGCTGCTTGGCGAAGACCCGG	175
QY	391	GTGGGGAAGTGGCTAAGTCTTGTGATGCTACTTTTGCCGGCGTTGACGGGAA	450
DB	176	GTGGCGACGTCGTACGACACGACATCGACCGCTGCGCCGCCGCTAGCCTGGAA	235
QY	451	CTTACACGACAGGGGAGACACAGGCGCGCCGCCACAGGGCT	494
DB	236	GATGCTCGCCGCCGCCGCGCACAGGCGCGCCGACAGAGGCT	279

Search completed: February 21, 2005, 09:01:39  
Job time : 4258 secs



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 145030**

**TO: Bao-Qun Li**  
**Location: rem/3a24/3c18**  
**Art Unit: 1648**  
**Wednesday, February 23, 2005**  
  
**Case Serial Number: 09/664363**

**From: Deirdre Arnold**  
**Location: Biotech-Chem Library**  
**REM 1A64**  
**Phone: 571-272-2532**

**Deirdre.Arnold@uspto.gov**

### **Search Notes**

*Please feel free to contact me if you have any questions or would like to amend the search.*

Thank you for using STIC services.

Regards,  
Deirdre Arnold

**This Page Blank (uspto)**

4

145030

mej

STIC-Biotech/ChemLib

From: Li, Bao-Qun  
Sent: Monday, February 14, 2005 8:40 AM  
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF SEQ ID NO; 22 IN APPLICATION SN.  
09,664,363. THANKS.  
Bao Qun Li M.D  
TC 1600  
Art Unit 1648  
Tel. 517-272-0904  
REM, 3C18  
Rm. 3D24

RECEIVED  
FEB 14 2005  
STIC-Biotech/ChemLib  
(STIC)

\*\*\*\*\*  
STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*  
Type of Search

NA Sequence: # \_\_\_\_\_  
AA Sequence :# \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*  
Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 09:05:38 : Search time 15711 Seconds  
(without alignments)  
11565.555 Million cell updates/sec

Title: US-09-664-363-22

Perfect score: 3750

Sequence: 1 TGGAGGCGGCTTTCACAGG.....ACTACCTCACCCGTCACCG 3750

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Database :

GenEmb1:\*  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_dl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sta:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	3750	100.0	3750	6 A28157	A28157 PT-NANBH mr
2	3750	100.0	3750	6 AR144051	AR144051 Sequence
3	3729.2	99.4	3750	6 A32203	A32203 NANBH PT P
4	3319.8	88.5	9359	14 AF313916	AF313916 Hepatitis
5	3312.2	88.3	9410	14 HPCK182	D50481 Hepatitis C
6	3302.6	88.1	9410	14 HPCK182	D50481 Hepatitis C
7	3297.8	87.9	9373	14 AF208024	AF208024 Hepatitis
8	3297.8	87.9	9573	14 AB049090	AB049090 Hepatitis
9	3294.6	87.9	9379	14 HCVPOLYP	AJ000009 Hepatitis
10	3293	87.8	9374	14 AF207753	AF207753 Hepatitis
11	3292.2	87.8	8642	6 AR488101	AR488101 Sequence
12	3292.2	87.8	8642	6 AX472292	AX472292 Sequence
13	3291.4	87.8	7989	6 AR406043	AR406043 Sequence
14	3291.4	87.8	7989	6 AX036255	AX036255 Sequence
15	3291.4	87.8	7989	12 SSR242652	AJ242652 Hepatitis
16	3291.4	87.8	8001	6 AR406042	AR406042 Sequence
17	3291.4	87.8	8001	6 AX036254	AX036254 Sequence
18	3291.4	87.8	8001	12 SSR242654	AJ242654 Hepatitis
19	3291.4	87.8	8637	6 AR406041	AR406041 Sequence

20	3291.4	87.8	8637	6 AX036253	AX036253 Sequence
21	3291.4	87.8	8637	12 SSR242651	AJ242651 Hepatitis
22	3291.4	87.8	8638	6 AR488121	AR488121 Sequence
23	3291.4	87.8	8638	6 AX472314	AX472314 Sequence
24	3291.4	87.8	8639	6 AR488100	AR488100 Sequence
25	3291.4	87.8	8639	6 AX472291	AX472291 Sequence
26	3291.4	87.8	8649	6 AR406044	AR406044 Sequence
27	3291.4	87.8	8649	6 AX036256	AX036256 Sequence
28	3291.4	87.8	8649	12 SSR242653	AJ242653 Hepatitis
29	3291.4	87.8	9605	6 AX739971	AX739971 Hepatitis
30	3291.4	87.8	9605	14 HC0238799	AJ238799 Hepatitis
31	3291.4	87.8	10690	6 AX739972	AX739972 Sequence
32	3291.4	87.8	11076	6 AR406040	AR406040 Sequence
33	3291.4	87.8	11076	6 AX036252	AX036252 Sequence
34	3289.8	87.7	12305	14 AP165055	AP165055 Hepatitis
35	3289.8	87.7	12305	6 AX937622	AX937622 Sequence
36	3289.8	87.7	12315	6 AX937621	AX937621 Sequence
37	3289	87.7	7065	6 A76577	AX937621 Sequence 9
38	3288.2	87.7	8638	6 AR488104	AR488104 Sequence
39	3288.2	87.7	8638	6 AR488105	AR488105 Sequence
40	3288.2	87.7	8638	6 AR488122	AR488122 Sequence
41	3288.2	87.7	8638	6 AX472296	AX472296 Sequence
42	3288.2	87.7	8638	6 AX472297	AX472297 Sequence
43	3288.2	87.7	8638	6 AX472315	AX472315 Sequence
44	3288.2	87.7	8643	6 AR488102	AR488102 Sequence
45	3288.2	87.7	8643	6 AX472294	AX472294 Sequence

## ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	CDS
A28157	A28157	PT-NANBH mRNA fragment from patent GB2239245.	A28157	A28157.1	GI:1248640	synthetic construct synthetic construct other sequences; artificial sequences. 1 (bases 1 to 3750)	Post-translational non-A non-B hepatitis viral polyproteins	Patent: GB 2239245-A 22-26-JUN-1991;				Location/Qualifiers	
												1. 3750	
												/organism="synthetic construct"	
												/mol_type="unassigned DNA"	
												/db_xref="taxon:32630"	
												/note="contig formed by cDNA clones from 3' end"	
												<1..>3750	
												/codon_start=1	
												/transl_table=11	
												/product="PT-NANBH polyprotein"	
												/protein_id="CAA01932.1"	
												/db_xref="GI:1248641"	
												/translation="MEGVFTGLTHDAHFLSOTKQAGNPFYLVAVQATVCARQAP	
												PSWDQMKCLIRLKPTRNGPTPLYLKGLAVQNEVLTITPTEFIMACMADELVST	
												WVLVGVLAALAAVCLTGSVIVGRILISGRPAIDRDELYQEDMEBSHLPY	
												IRGQMLAOPFOKMLGLLOTAKTAAEAAPVESRMALRETFMAAMNFTSGIYL	
												AGLSTTPGNPAIASLMAPTASVTSPLTSTOSTLITLIGCVAAQLAPAAASPVAG	
												IKGAAGSTIGLGVLDIAGVAGVAGALVAFKWSGEMPTEDLVNLTALISFGA	
												LIVGVVCAALIRRHVGGGAVQMMNRLLAFASRGNHVPRTYVPSDAARTQILS	
												DLITDILRLRQWINDSTSCSGSLMDVDMICTVLADEKTKMQLPRLPVP	
												FTSCQGVKVRGDIQMTSCGQITIGHVNGSMRLVGCRTCSNMHGFPIYAV	
												TVGQCTSPAPVSRALMVAEVEYVTRVGVDPHYVTSMTDNVKPCQVAPBEFT	
												EVDGVLRHVPARCKRLAREYTFQVGNQYVGSQLPCEPDDVAVLTSMITPSEHI	
												THETAKRLARSGPPRLASSASQLSAPESKATYITONDPPADILFANLIRHENG	
												DLTRVESNKKVILDSFDPLRAEBDEREVSAEILIRKSKTTPPAVAPARPDYNPL	
												LSWKAIPDVPPVHGCPLPPTKPTPIPPRRKRTVLTSTVSSALALATKAFSS	
												BSAVDSGTATAPDPQSDGAGSVASVSSMPLEGGPCDPLSDGSMSTVSEBAG	
												EDVVCSSMYTWTGALITPCAAESKLPINALNSILRRHNNVYATTSASQROKVK	

TFDQLQJLBHJYDVLKEMAKASITYKALUSVEAEKCUJTPHSAKSFEYGAQDVN  
 LSKA INHISVWDLDETEJDPITIAKOREFVOPPERGKRPDALIYFPUJGSR  
 VCEKARADIVSWLTPQWMSSEGVOPSEKORVEFIVNMAKSRTKPFADVPCPS  
 TVTENDI RVESVIYOCDDLAPEAROLAIRLITRELYIYGEPLNSKONGYBRCA  
 LTTSCGNUTITREKASACPAKALQDTCMLVCGDGLIYVCSAGJONGBASIVFTFA  
 TTRYSAPGCPDPEYDLTEITSCSNVSNVADHASKRYYVLTBP"

ORIGIN

Query Match	100.0%;	Score 3750;	DB 6;	Length 3750;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3750;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	1	GGGAGAGGGGCTTTACAGAGGCTCAACCAACGGAGATGCCCACTTCCTGCCCCAACAAAG	60
Db	1	1	TGGGAGGGGGCTTTCACAGAGCTCAACCAACGGAGATGCCCACTTCCTGCCCCAACAAAG	60
QY	61	61	CAGGACGAGAGCAACATTCCCCTACCTGGTGGGGGTACCAAGGCTACTGTGCGCTAGGGCC	120
Db	61	61	CAGGACGAGAGCAACATTCCCCTACCTGGTGGGGGTACCAAGGCTACTGTGTGGCGTAGGGCC	120
QY	121	121	CAGGCCCCACCTTCATCATATGGGATCAAAATGTGAAAGTGTCTCATACGGCTAAAGCTACT	180
Db	121	121	CAGGCCCCACCTTCATCATATGGGATCAAAATGTGAAAGTGTCTCATACGGCTAAAGCTACT	180
QY	181	181	CTGGCGGGGCGCAACACCTTGTCTGTATATAGGCTTGGGAGCCGCTCAAAAAGAGGTCAACCTC	240
Db	181	181	CTGGCGGGGCGCAACACCTTGTCTGTATATAGGCTTGGGAGCCGCTCAAAAAGAGGTCAACCTC	240
QY	241	241	ACACACCCCATATACCAAAATTCATCATATGGGATCATGTACGGCACTGGAGGGTGGTCAAG	300
Db	241	241	ACACACCCCATATACCAAAATTCATCATATGGGATCATGTACGGCACTGGAGGGTGGTCAAG	300
QY	301	301	AGCACCTGGGTGCTGGTGGGCGGGGTCTTTGCAAGCTTGTGCTGTAGCTATTTGCTTGA	360
Db	301	301	AGCACCTGGGTGCTGGTGGGCGGGGTCTTTGCAAGCTTGTGCTGTAGCTATTTGCTTGA	360
QY	361	361	GGCAGCGTGTCTATTGTGGGTATGATCATCTTGTCCGGCGGCGCGCTATTGTTCCTGAG	420
Db	361	361	GGCAGCGTGTCTATTGTGGGTATGATCATCTTGTCCGGCGGCGCGCTATTGTTCCTGAG	420
QY	421	421	AGGGAAGCTCTACACAGAGTTGCATATGATATGAAAGTGGGCGTGCACCTCCCTTAC	480
Db	421	421	AGGGAAGCTCTCTACACAGAGTTGCATATGATATGAAAGTGGGCGTGCACCTCCCTTAC	480
QY	481	481	ATCGAGCAGGGAATGAGCTTCGCCGAGCAGTTCAAGCAAAAACGCTCGGTTTGTGTGAG	540
Db	481	481	ATCGAGCAGGGAATGAGCTTCGCCGAGCAGTTCAAGCAAAAACGCTCGGTTTGTGTGAG	540
QY	541	541	ACAGCCACCAAGACACGAGGCGCTGCTCCCTGTGTGTAGATCCAATGTGGCAGCCCTT	600
Db	541	541	ACAGCCACCAAGACACGAGGCGCTGCTCCCTGTGTGTAGATCCAATGTGGCAGCCCTT	600
QY	601	601	GAGACCTTCTGGGCGCAACACATATGTGAACCTTATCAGCGGGATTCAGTACTTACAGGC	660
Db	601	601	GAGACCTTCTGGGCGCAACACATATGTGAACCTTATCAGCGGGATTCAGTACTTACAGGC	660
QY	661	661	TTGTCCACTCTGCTGTGGGAATCCCGGATTTGCATCACTATAGGCTTCAACGCTCTGTC	720
Db	661	661	TTGTCCACTCTGCTGTGGGAATCCCGGATTTGCATCACTATAGGCTTCAACGCTCTGTC	720
QY	721	721	ACTAGCCCGCTACCAACCCCAATCTACCCCTTGTAAATCTCTGTGGGGATGGGTAGCC	780
Db	721	721	ACTAGCCCGCTACCAACCCCAATCTACCCCTTGTAAATCTCTGTGGGGATGGGTAGCC	780
QY	781	781	GGCCCACTGTCTCCCCCAGTGTCTGTTCAAGCTTTGTATAGGCGCGGCAATTTGCTGTGCG	840
Db	781	781	GGCCCACTGTCTCCCCCAGTGTCTGTTCAAGCTTTGTATAGGCGCGGCAATTTGCTGTGCG	840
QY	841	841	GCTGTTGGCAGATATAGGCTTGTGGAAAGTGTGTGGAATCTGTGGCGGCTATATGAGCA	900
Db	841	841	GCTGTTGGCAGATATAGGCTTGTGGAAAGTGTGTGGAATCTGTGGCGGCTATATGAGCA	900
QY	901	901	GGAGTGGCAGAGCGCTGTGGCTTTTAAAGTATATAGCGCGCAATGCTTCCACGAG	960

Db	901	GGAETGGCAGGCGGGCTGTGTGGCTTTTAAGTATGAGCGGCAAAATGCCCTCCAGCGAG	960
Qy	961	GACCTGGTTAACTTACTTCCCTGCGCATCTCTCTCTGTGTGCCCTGTGTGTGTGGGGTCTGT	1020
Db	961	GACCTGGTTAACTTACTTCCCTGCGCATCTCTCTCTGTGTGCCCTGTGTGTGTGGGGTCTGT	1020
Qy	1021	TGGCCAGAGAACTGCGCGTCCGCGCAGTGGGTCCAGGGGAGGGGGGCTGTGCAGTGGATGAAC	1080
Db	1021	TGGCCAGAGAACTGCGCGTCCGCGCAGTGGGTCCAGGGGAGGGGGGCTGTGCAGTGGATGAAC	1080
Qy	1081	CGGCTGATAGCGTTGCGCTCGCGGGGTAAACCATTTTCCCGCAGCACTATGTGTCCAGAG	1140
Db	1081	CGGCTGATAGCGTTGCGCTCGCGGGGTAAACCATTTTCCCGCAGCACTATGTGTCCAGAG	1140
Qy	1141	AGCGACGCGGAGCAGCGTGTCACTCAATCTCTTCGACCTTATCATCCCAACTGTGTG	1200
Db	1141	AGCGACGCGGAGCAGCGTGTCACTCAATCTCTTCGACCTTATCATCCCAACTGTGTG	1200
Qy	1201	AAGAGGCTCCACCACTGTGATTAACGAGGACCTGTCCACGCCCTGTCTCCGGTCTGTGGCTA	1260
Db	1201	AAGAGGCTCCACCACTGTGATTAACGAGGACCTGTCCACGCCCTGTCTCCGGTCTGTGGCTA	1260
Qy	1261	AGGATGTGTTGGGACTGGATATGACAGTTTGTGGCTGACTTCAAGACTGGAGCTGCAGTTC	1320
Db	1261	AGGATGTGTTGGGACTGGATATGACAGTTTGTGGCTGACTTCAAGACTGGAGCTGCAGTTC	1320
Qy	1321	AAGCTCTGCGCGCGATTAACCGGAGTCCCTTTTTCATATGCCAACGCTGGGTACAGAGGG	1380
Db	1321	AAGCTCTGCGCGCGATTAACCGGAGTCCCTTTTTCATATGCCAACGCTGGGTACAGAGGG	1380
Qy	1381	GCTGTGGCGGGAGACCGGATCATGTACAGACCCTGTCTCATGTGGAGACACAGATACCGGA	1440
Db	1381	GCTGTGGCGGGAGACCGGATCATGTACAGACCCTGTCTCATGTGGAGACACAGATACCGGA	1440
Qy	1441	CATGTCAAAAACGGTTCCATGAGGATGTTGGGCTTAAGACCTGTAGTAACATGTGGCAT	1500
Db	1441	CATGTCAAAAACGGTTCCATGAGGATGTTGGGCTTAAGACCTGTAGTAACATGTGGCAT	1500
Qy	1501	GGAACATTTCCCATCAACGCATACACCAACCGGGCCCTGTGACGCCCTCCCGACGCCCAAC	1560
Db	1501	GGAACATTTCCCATCAACGCATACACCAACCGGGCCCTGTGACGCCCTCCCGACGCCCAAC	1560
Qy	1561	TATTTACAGGGCGCTGTGTGGCGGGTGTGCTGTGAGAGATACGTGAGGTTTACCGGGGTGGG	1620
Db	1561	TATTTACAGGGCGCTGTGTGGCGGGTGTGCTGTGAGAGATACGTGAGGTTTACCGGGGTGGG	1620
Qy	1621	GATTTTCCACTACGTGACGAGACATCACTGACCAACGTAACAAATCCCGTGCAGGTTTCCA	1680
Db	1621	GATTTTCCACTACGTGACGAGACATCACTGACCAACGTAACAAATCCCGTGCAGGTTTCCA	1680
Qy	1681	GCCCCGGAATTTCTTACAGAAAGTGGATGGGGTGTGGGCTGTGACAGGTAACGCTCCGGCGTGC	1740
Db	1681	GCCCCGGAATTTCTTACAGAAAGTGGATGGGGTGTGGGCTGTGACAGGTAACGCTCCGGCGTGC	1740
Qy	1741	AAACCTCTCTACCGGAGAGGATCAATTCAGGTCGGGCTCAACCAATACCTGTGTGGG	1800
Db	1741	AAACCTCTCTACCGGAGAGGATCAATTCAGGTCGGGCTCAACCAATACCTGTGTGGG	1800
Qy	1801	TGCGAGCTTCCATGTGAGACCCGGAACCGGATGTACAGATGCTCACTTCATGTCTACCGAC	1860
Db	1801	TGCGAGCTTCCATGTGAGACCCGGAACCGGATGTACAGATGCTCACTTCATGTCTACCGAC	1860
Qy	1861	CCCTCCCAATACACGACGAGACGAGCTTAAGGCGAGGCTGTGGCAGGGGATCTCCCCCTCC	1920
Db	1861	CCCTCCCAATACACGACGAGACGAGCTTAAGGCGAGGCTGTGGCAGGGGATCTCCCCCTCC	1920
Qy	1921	TTGGCCAGACTTTCAGCTAGCCAGTTGTCTGTGCGCTTCTTGAAAGGCGACATACATTAAC	1980
Db	1921	TTGGCCAGACTTTCAGCTAGCCAGTTGTCTGTGCGCTTCTTGAAAGGCGACATACATTAAC	1980
Qy	1981	CAAAATGACTTCCAGACGCTGACCTCATGAGGCGCAACTTCTGTGTGGCGGCATGAGATG	2040
Db	1981	CAAAATGACTTCCAGACGCTGACCTCATGAGGCGCAACTTCTGTGTGGCGGCATGAGATG	2040





Db 1 TGGAGAGGCGCTCTTCAAGGCTCAACCACGTGATGCCCACTTCTGTCCTCCAAACAAAG 60  
QY 61 CAGGACGAGAGCAACTTCCCTCACTGGTGGCGTACAGGCTACTGTTGTCGCTAGAGGCC 120  
Db 61 CAGGACGAGAGCAACTTCCCTCACTGGTGGCGTACAGGCTACTGTTGTCGCTAGAGGCC 120  
QY 121 CAGGCCCACTTCATCATGAGATCAAAATGTGAAAGTGTCTCATACGCTTAAAGCTTACT 180  
Db 121 CAGGCCCACTTCATCATGAGATCAAAATGTGAAAGTGTCTCATACGCTTAAAGCTTACT 180  
QY 181 CTGCGGGGCGCAACACCTTGTGTATAGGCTGGAGGCCCTCCAAAACGAGGTCACTTC 240  
Db 181 CTGCGGGGCGCAACACCTTGTGTATAGGCTGGAGGCCCTCCAAAACGAGGTCACTTC 240  
QY 241 ACACACCCCAATAACCAATTCAATGAGCATGATGACCGACCTGGAGGTGTCACG 300  
Db 241 ACACACCCCAATAACCAATTCAATGAGCATGATGACCGACCTGGAGGTGTCACG 300  
QY 301 AGCACCCTGGTGTGTGGGCGGGGTCCTTTCAGACTCTGGCTGCTATTTGCTTGAACA 360  
Db 301 AGCACCCTGGTGTGTGGGCGGGGTCCTTTCAGACTCTGGCTGCTATTTGCTTGAACA 360  
QY 361 GGCAGCGTGTCTATTTGGGTAGATATCTTGTCCGGGCGCGGCTATTTGCTTCCGAC 420  
Db 361 GGCAGCGTGTCTATTTGGGTAGATATCTTGTCCGGGCGCGGCTATTTGCTTCCGAC 420  
QY 421 AGGGAAGTCTCTACAGAGATTGATGATGGAAGAGTGGCGGTGACCTCCCTTAC 480  
Db 421 AGGGAAGTCTCTACAGAGATTGATGATGGAAGAGTGGCGGTGACCTCCCTTAC 480  
QY 481 ATGAGACAGAGAAATGACAGTCCGCGAGCACTTCAAGCAAAAGCGCTCGGTTGCTGAC 540  
Db 481 ATGAGACAGAGAAATGACAGTCCGCGAGCACTTCAAGCAAAAGCGCTCGGTTGCTGAC 540  
QY 541 ACAGCCCAACAGAGAGGAGGCGGCTGCTCCGCTGGTGAATCAAGTGGCGAGCCCTT 600  
Db 541 ACAGCCCAACAGAGAGGAGGCGGCTGCTCCGCTGGTGAATCAAGTGGCGAGCCCTT 600  
QY 601 GAGACCTTCTGGGCGAAACATGTGGAATCTTCAATGCGGAGATACGATCTTACAGAGC 660  
Db 601 GAGACCTTCTGGGCGAAACATGTGGAATCTTCAATGCGGAGATACGATCTTACAGAGC 660  
QY 661 TTGTCACTCTGCTGGGAAATCCGGATGTCATCACTGATGGGTTCAACAGCTCTGTC 720  
Db 661 TTGTCACTCTGCTGGGAAATCCGGATGTCATCACTGATGGGTTCAACAGCTCTGTC 720  
QY 721 ACTAGCCCGCTCACACCAATCTACCTCTGCTTAAATCTCTGGGGGATGGGTAGCC 780  
Db 721 ACTAGCCCGCTCACACCAATCTACCTCTGCTTAAATCTCTGGGGGATGGGTAGCC 780  
QY 781 GCCCAACTGCTCCCGCAGTGTCTTCAAGCTTGTAGGGCGCGCATTTGCTGTGCG 840  
Db 781 GCCCAACTGCTCCCGCAGTGTCTTCAAGCTTGTAGGGCGCGCATTTGCTGTGCG 840  
QY 841 GCTGTGTGAGCAATAGGCTTGGAGAGTGTGTGACATCTTGGCGGCTATGAGACA 900  
Db 841 GCTGTGTGAGCAATAGGCTTGGAGAGTGTGTGACATCTTGGCGGCTATGAGACA 900  
QY 901 GAGTGTGAGAGCGGCTGTGGCTTTAAGTCTATGACGCGCAATATGCTTCCACCGAG 960  
Db 901 GAGTGTGAGAGCGGCTGTGGCTTTAAGTCTATGACGCGCAATATGCTTCCACCGAG 960  
QY 961 GACCTGTAACTTACTCCCGGCAATCTCTCTCTGTGTGCTGTGTGTGTGTGTGTGT 1020  
Db 961 GACCTGTAACTTACTCCCGGCAATCTCTCTCTGTGTGCTGTGTGTGTGTGTGTGT 1020  
QY 1021 TGGCAGAGCAATCTGCTGCGCAGTGGGTCAAGGAGGAGGAGCTGTGAGTGAATAC 1080  
Db 1021 TGGCAGAGCAATCTGCTGCGCAGTGGGTCAAGGAGGAGGAGCTGTGAGTGAATAC 1080  
QY 1081 CGGCTGATAGCGTTGCGCTCGCGGGGTAAACATGTTCCCGCAGCACTATGTGCGAG 1140  
Db 1081 CGGCTGATAGCGTTGCGCTCGCGGGGTAAACATGTTCCCGCAGCACTATGTGCGAG 1140

QY 1141 AGCGAGCCGAGACAGTGTCACTAGATCTCTCGACCTTACTATCAACCACTGTG 1200  
Db 1141 AGCGAGCCGAGACAGTGTCACTAGATCTCTCGACCTTACTATCAACCACTGTG 1200  
QY 1201 AAGAGCTCCACCAATGATTTAAACAGAGCTGTCCAGCCCTGCTCCGGCTGTGGCTA 1260  
Db 1201 AAGAGCTCCACCAATGATTTAAACAGAGCTGTCTCCAGCCCTGCTCCGGCTGTGGCTA 1260  
QY 1261 AAGGATGTTGGAGCTGATATGACAGTGTGGCTTCAAGACTGGCTCCAGTCC 1320  
Db 1261 AAGGATGTTGGAGCTGATATGACAGTGTGGCTTCAAGACTGGCTCCAGTCC 1320  
QY 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCAACGTGGGTAAAGAGG 1380  
Db 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCAACGTGGGTAAAGAGG 1380  
QY 1381 GTCTGGCGGGAGACCGCATCATGACACCACTGCTCATGTGAGCAAGATCACCGGA 1440  
Db 1381 GTCTGGCGGGAGACCGCATCATGACACCACTGCTCATGTGAGCAAGATCACCGGA 1440  
QY 1441 CATGTCAAAAAGGTTCCATGAGGATGTTGGGCTTAAGACCTGTATACATGTGGCAT 1500  
Db 1441 CATGTCAAAAAGGTTCCATGAGGATGTTGGGCTTAAGACCTGTATACATGTGGCAT 1500  
QY 1501 GGAACATTCGCCATCAACGATATCAACAGGCGCCCTGACGCGCTCCAGCGCCAAAC 1560  
Db 1501 GGAACATTCGCCATCAACGATATCAACAGGCGCCCTGACGCGCTCCAGCGCCAAAC 1560  
QY 1561 TATTCAGAGCGCTGTGGCGGGTGTCTGAGAGATAGTGAAGTTACGCGGTTGGG 1620  
Db 1561 TATTCAGAGCGCTGTGGCGGGTGTCTGAGAGATAGTGAAGTTACGCGGTTGGG 1620  
QY 1621 GATTTCACTAGTGAAGCATATGACATGACATGATGATGATGATGATGATGATGATG 1680  
Db 1621 GATTTCACTAGTGAAGCATATGACATGACATGATGATGATGATGATGATGATGATG 1680  
QY 1681 GCGCCCGAATTCCTTCAACAGAGTGTGGGTTGCGGCTGACAGATGCTCCGCGGTGC 1740  
Db 1681 GCGCCCGAATTCCTTCAACAGAGTGTGGGTTGCGGCTGACAGATGCTCCGCGGTGC 1740  
QY 1741 AAACCTCTCTACGAGAGAGGTCAATTCAGATGAGTGTGAGTGTGAGTGTGAGTGT 1800  
Db 1741 AAACCTCTCTACGAGAGAGGTCAATTCAGATGAGTGTGAGTGTGAGTGTGAGTGT 1800  
QY 1801 TCGCACTCCCATGAGAGCGGCGGAAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1860  
Db 1801 TCGCACTCCCATGAGAGCGGCGGAAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1860  
QY 1861 CCCTCCCATGAGAGCGGCGGAAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1920  
Db 1861 CCCTCCCATGAGAGCGGCGGAAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1920  
QY 1921 TTGGCAGCTCTTCACTAGACCATGTTGTCTGCGCTTCTCGAAGCGACATCATTAAC 1980  
Db 1921 TTGGCAGCTCTTCACTAGACCATGTTGTCTGCGCTTCTCGAAGCGACATCATTAAC 1980  
QY 1981 CAAAATGACTTCCCAACGCTGACCTCATGAGGCAACCTCTGAGGAGGAGTGAATG 2040  
Db 1981 CAAAATGACTTCCCAACGCTGACCTCATGAGGCAACCTCTGAGGAGGAGTGAATG 2040  
QY 2041 GCGGAGGAGATTAACCGCGTGAAGTCAAGAAACAGATGATATCTTTCGAC 2100  
Db 2041 GCGGAGGAGATTAACCGCGTGAAGTCAAGAAACAGATGATATCTTTCGAC 2100  
QY 2101 CGGCTCGAGCGGAGAGATGAGCGGAAATGTCCGTCCCGCGGAGATCTTCGCGAAA 2160  
Db 2101 CGGCTCGAGCGGAGAGATGAGCGGAAATGTCCGTCCCGCGGAGATCTTCGCGAAA 2160  
QY 2161 TCCAAAGAAATTCACCAAGAGATCCCGCATGGGCAAGCCGGAATTAACAACCTCCGCTG 2220  
Db 2161 TCCAAAGAAATTCACCAAGAGATCCCGCATGGGCAAGCCGGAATTAACAACCTCCGCTG 2220

QY	2221	CTGGAGTCTCTGGAAAGCCCGGAGACTACGCTCCCTCAGGAGTAACTAGGGTGGCCACTGGCA	2280
Db	2221	CTGGAGTCTCTGGAAAGCCCGGAGCTTAGGTCCTCTCAGTGTGACATAGGGTGGCCACTGGCA	2280
QY	2281	CCTACTAAGACCCCTCTATACCACTCCACGAGAGAAAGAGACAGTGTCTGACAGAA	2340
Db	2281	CCTACTAAGACCCCTCTATACCACTCCACGAGAGAAAGAGACAGTGTCTGACAGAA	2340
QY	2341	TCACACGTGTCTTCTGCCCCCTGGCGGAGCTTGCACAAAGGCTTTCGTAGCTCCGAAACCG	2400
Db	2341	TCACACGTGTCTTCTGCCCCCTGGCGGAGCTTGCACAAAGGCTTTCGTAGCTCCGAAACCG	2400
QY	2401	TCGGCCGTCGACAGGGGACCGGCAACCGCCCTCTGTACCAACCTCCGACGAGCGGCGGA	2460
Db	2401	TCGGCCGTCGACAGGGGACCGGCAACCGCCCTCTGTACCAACCTCCGACGAGCGGCGGA	2460
QY	2461	GCAGAGTCTGACGTTGAGTCCGTAATCTCTCAATGCCCTTCTGAGGGGAGACCGGGGAGAC	2520
Db	2461	GCAGAGTCTGACGTTGAGTCCGTAATCTCTCAATGCCCTTCTGAGGGGAGACCGGGGAGAC	2520
QY	2521	CCCGATCTCAGCGACGGGTCCTTGATCTTACCGTAGTAGAGAGCCGTTAGAGACGTCGTC	2580
Db	2521	CCCGATCTCAGCGACGGGTCCTTGATCTTACCGTAGTAGAGAGCCGTTAGAGACGTCGTC	2580
QY	2581	TGCTGCTCGATGTCTTACACATAGACAGGCGTCTGATCACGCGCATGCGCTGCGGAGAA	2640
Db	2581	TGCTGCTCGATGTCTTACACATAGACAGGCGTCTGATCACGCGCATGCGCTGCGGAGAA	2640
QY	2641	AGCAAGCTGSCCATCAACGCGCTTGAGCAACTCTTGTGCTGCGTCAACAACATGGTCTAC	2700
Db	2641	AGCAAGCTGSCCATCAACGCGCTTGAGCAACTCTTGTGCTGCGTCAACAACATGGTCTAC	2700
QY	2701	GCTACCAATCCCGGAGCGCAAGCCGAGCGGAGAAAGAGTCAACCTTTGACAGACTGGCA	2760
Db	2701	GCTACCAATCCCGGAGCGCAAGCCGAGCGGAGAAAGAGTCAACCTTTGACAGACTGGCA	2760
QY	2761	ATCTCTGACGATCACTACAGGACGTGTCTACAGAGATAGAAAGGAGAGCGTCCACAGTT	2820
Db	2761	ATCTCTGACGATCACTACAGGACGTGTCTACAGAGATAGAAAGGAGAGCGTCCACAGTT	2820
QY	2821	AAGGTAAAGCTCTATCAGTAGAGAGACCTGCAAGCTGACGAGCCGCCACATTCGGGCCAA	2880
Db	2821	AAGGTAAAGCTCTATCAGTAGAGAGACCTGCAAGCTGACGAGCCGCCACATTCGGGCCAA	2880
QY	2881	TCTAAATTTGGCTATGGGGCAAAAGGACGTCGGAACCTATCAGCAAGGCCATTTAACAC	2940
Db	2881	TCTAAATTTGGCTATGGGGCAAAAGGACGTCGGAACCTATCAGCAAGGCCATTTAACAC	2940
QY	2941	ATCCGCTCCGTGTGGGAGGAACTTGTGGAAAGCACTGAAACCAATTTGACACCACTATC	3000
Db	2941	ATCCGCTCCGTGTGGGAGGAACTTGTGGAAAGCACTGAAACCAATTTGACACCACTATC	3000
QY	3001	ATGGCAAAAATAGAGTTTCTGCGGTCCAACAGAGAGAGAGCGGCAACCGCACTGCGC	3060
Db	3001	ATGGCAAAAATAGAGTTTCTGCGGTCCAACAGAGAGAGAGCGGCAACCGCACTGCGC	3060
QY	3061	CTTATCGTGTCCCAAGCTTGGGGGTCCGCTGTGTGCGAGAAATAGGACCTCTATGACGTG	3120
Db	3061	CTTATCGTGTCCCAAGCTTGGGGGTCCGCTGTGTGCGAGAAATAGGACCTCTATGACGTG	3120
QY	3121	GTCTCCACCTCCCTCAGGCTGTGATGGGCTCTCTCTGACGGAATTCAGTATTTCTCTGG	3180
Db	3121	GTCTCCACCTCCCTCAGGCTGTGATGGGCTCTCTCTGACGGAATTCAGTATTTCTCTGG	3180
QY	3181	CAGCGGGTCAGGTTCTCTGTGAAAGCCCTGGAATCAAAAGAAACCCCTATGCGCTTTGCA	3240
Db	3181	CAGCGGGTCAGGTTCTCTGTGAAAGCCCTGGAATCAAAAGAAACCCCTATGCGCTTTGCA	3240
QY	3241	TATGACACCCGCTGTTTGACTCAACAGTCACTGGAATGACATCCGTGTAGAGAGTCA	3300
Db	3241	TATGACACCCGCTGTTTGACTCAACAGTCACTGGAATGACATCCGTGTAGAGAGTCA	3300
QY	3301	ATTATCAATGTTGACTTGCGCCCGAAGCCAGACGAGCCCATAGAGTCCCTCAACAG	3360

[illegible]

## ORIGIN

VEKALYDVNSTLPOAVMGSSVYQYPSQGVPEPLVNAWKSKEPMGFAYDRCPDS  
TTEENDI RVEESIYCCDLAPEARQAI RSLRILYIGPLTNSGONCGRRPRRAGV  
LTSGNLTLCVYKASACBAKLODCTMLVCGDGLVVICESHTQEDBASLKFVFEA  
MTRYSAPDPDPPPEYDELITSCSSNSVSAHDA SGRVYLLTRDP"

Query Match	99.4%; Score 3729.2; DB 6; Length 3750;	
Best Local Similarity	99.7%; Pred. No. 0;	
Matches 3737; Conservative	0; Mismatches 13; Indels 0; Gaps 0;	
QY	1	TEGGAGGGCGCTTTTCACAGGCTTCACCACGTGGATGCCACTTCTGTCCCAACAAAG 60
DB	1	TGGAGGGCGCTTTCACAGGCTTCACCACGTGGATGCCACTTCTGTCCCAACAAAG 60
QY	61	CAGGACGAGAGCAACTTCCCTTACTGTGGTGGCGTACAGGCTACTGTGTGGCGTGGAGGCC 120
DB	61	CAGGACGAGAGCAACTTCCCTTACTGTGGTGGCGTACAGGCTACTGTGTGGCGTGGAGGCC 120
QY	121	CAGGCCCCACTTCATCATGTGGATCAAAATGTGAAAGTGTCTCATACGGCTTAAAGCCTACT 180
DB	121	CAGGCCCCACTTCATCATGTGGATCAAAATGTGAAAGTGTCTCATACGGCTTAAAGCCTACT 180
QY	181	CTGCGGGGCGCAACACCTTGTGTATAGCTGGAGCGCTCCAAAACAGATCACCTTC 240
DB	181	CTGCGGGGCGCAACACCTTGTGTATAGCTGGAGCGCTCCAAAACAGATCACCTTC 240
QY	241	ACACACCCCATTAACAAATTCATCATGTGGATGCAATGTCAGCCGACCTGGAGGTCTGACG 300
DB	241	ACACACCCCATTAACAAATTCATCATGTGGATGCAATGTCAGCCGACCTGGAGGTCTGACG 300
QY	301	AGCACCTGGTGTGTGGGCGGGGTCTTGACAGCTTGTGCTGTATTTGCTTGACAAACA 360
DB	301	AGCACCTGGTGTGTGGGCGGGGTCTTGACAGCTTGTGCTGTATTTGCTTGACAAACA 360
QY	361	GGCAGGCTGTCTATTGGGTAGGATCATCTTGTCCGGGCGGGCTATTTGTTCCCGAC 420
DB	361	GGCAGGCTGTCTATTGGGTAGGATCATCTTGTCCGGGCGGGCTATTTGTTCCCGAC 420
QY	421	AGGGAAGTCTCTACAGAGATTGATGAGATGGAAGTGGCGGTGCGACTCCCTTAC 480
DB	421	AGGGAAGTCTCTCTACAGAGATTGATGAGATGGAAGTGGCGGTGCGACTCCCTTAC 480
QY	481	ATCGAGCAGGAAATGCACTGCGCGAGCAGTTCAAGCAAAAAGCGCTCGGGTTCGTCAG 540
DB	481	ATCGAGCAGGAAATGCACTGCGCGAGCAGTTCAAGCAAAAAGCGCTCGGGTTCGTCAG 540
QY	541	ACAGCCACCAAGCAGAGGCGCGCTCTCCGTGTGAGATCCCAATGTGGCGAGCCCTT 600
DB	541	ACAGCCACCAAGCAGAGGCGCGCTCTCCGTGTGAGATCCCAATGTGGCGAGCCCTT 600
QY	601	GAGACCTTCTGGGGGAAACATGTGGAATTCATCAGCGGGAATACGTAATTAGCAGGC 660
DB	601	GAGACCTTCTGGGGGAAACATGTGGAATTCATCAGCGGGAATACGTAATTAGCAGGC 660
QY	661	TTGTCCAATCTGCTGGGAAATCCCGCATGTGATCACTGATGGGTTCAACAGCCTGTCT 720
DB	661	TTGTCCAATCTGCTGGGAAATCCCGCATGTGATCACTGATGGGTTCAACAGCCTGTCT 720
QY	721	ACTAGCCCCGTCAACCACTCAATCACTCTGTCTTAAATCTCTGGGGGGAATGGGTGCC 780
DB	721	ACTAGCCCCGTCAACCACTCAATCACTCTGTCTTAAATCTCTGGGGGGAATGGGTGCC 780
QY	781	GCCCAACTCGCTCCCGCAGTGTCTTCACTTGTAGGCGCGGCAATTTGCTGGTGGC 840
DB	781	GCCCAACTCGCTCCCGCAGTGTCTTCACTTGTAGGCGCGGCAATTTGCTGGTGGC 840
QY	841	GCTGTGGCAGATAGGCTTTGGGAGGTGCTTGAATCACTTGGCGGGGTATAGAGCA 900
DB	841	GCTGTGGCAGATAGGCTTTGGGAGGTGCTTGAATCACTTGGCGGGGTATAGAGCA 900
QY	901	GAGTGGCAGGCGGCTGTGGCTTTAAGTCAATGAGCGCGAATGCCCTCCACCGAG 960
DB	901	GAGTGGCAGGCGGCTGTGGCTTTAAGTCAATGAGCGCGAATGCCCTCCACCGAG 960

QY	961	GACCTGGTTAACTTAATCCCTGGCAATCTCTCTCTGTGGTCCCTGTGCTGGGGTCTGTG 1020
DB	961	GACCTGGTTAACTTAATCCCTGGCAATCTCTCTCTGTGGTCCCTGTGCTGGGGTCTGTG 1020
QY	1021	TGCGCAGCATATCTGCTGTGGCAGCTGGGTCCAGGGGAGGGGCTGTCAATGTGATGAAC 1080
DB	1021	TGCGCAGCATATCTGCTGTGGCAGCTGGGTCCAGGGGAGGGGCTGTGATGTGATGAAC 1080
QY	1081	CGGCTGATAGCGTGTGGCTCGGGGGTAAACAATGTTTCCCAACGCACTATGTGCCAGAG 1140
DB	1081	CGGCTGATAGCGTGTGGCTCGGGGGTAAACAATGTTTCCCAACGCACTATGTGCCAGAG 1140
QY	1141	AGCGACCGCGCAGCAGGTGTCACTCAGATCTCTCCGACCTTAACTATCAACCAACTGTGG 1200
DB	1141	AGCGACCGCGCAGCAGGTGTCACTCAGATCTCTCCGACCTTAACTATCAACCAACTGTGG 1200
QY	1201	AAGAGCTCCACCAAGTGAATTAACAGAGACTGTCCAAGCCCTGTCTCCGGCTGTGGCTTA 1260
DB	1201	AAGAGCTCCACCAAGTGAATTAACAGAGACTGTCCAAGCCCTGTCTCCGGCTGTGGCTTA 1260
QY	1261	AGGAGTGTGGGACTGGAATATGCAAGTTTGGCTGAATTAAGCACTGGCTCCAGTCC 1320
DB	1261	AGGAGTGTGGGACTGGAATATGCAAGTTTGGCTGAATTAAGCACTGGCTCCAGTCC 1320
QY	1321	AAGCTCTGCGCGATTAACCGGGAGTCCCTTTTCTCATATGCCAACGTGGGTAACAAGGGG 1380
DB	1321	AAGCTCTGCGCGATTAACCGGGAGTCCCTTTTCTCATATGCCAACGTGGGTAACAAGGGG 1380
QY	1381	GTTGCGGGGAGAGCGGCATCATGACACCACTGCTCATGTGAGACACAGATCACCGGA 1440
DB	1381	GTTGCGGGGAGAGCGGCATCATGACACCACTGCTCATGTGAGACACAGATCACCGGA 1440
QY	1441	CATGTCAAAAAGCGGTTCAATGAGGATGTTGGGCTTAAGACCTGTATGTAGACACAGATCA 1500
DB	1441	CATGTCAAAAAGCGGTTCAATGAGGATGTTGGGCTTAAGACCTGTATGTAGACACAGATCA 1500
QY	1501	GGAACTTCCCATCAACGATACACACGAGGCGCTTGAACGCGCTTCCCAAGCGCAAAAC 1560
DB	1501	GGAACTTCCCATCAACGATACACACGAGGCGCTTGAACGCGCTTCCCAAGCGCAAAAC 1560
QY	1561	TATTCAGGGCGCTGTGGGGGTGTGCTGAGAGATGATGAGAGTTACGCGGGGTGGGG 1620
DB	1561	TATTCAGGGCGCTGTGGGGGTGTGCTGAGAGATGATGAGAGTTACGCGGGGTGGGG 1620
QY	1621	GATTTCCATGATGATGAGATGACCACTGACCAAGTAAATGCCCGTGCAGATTCCA 1680
DB	1621	GATTTCCATGATGATGAGATGACCACTGACCAAGTAAATGCCCGTGCAGATTCCA 1680
QY	1681	GCCCCGGAATCTTCAACAGAGTGAATGGGATGCGGCTGACAGATACGCTCGGCGTGC 1740
DB	1681	GCCCCGGAATCTTCAACAGAGTGAATGGGATGCGGCTGACAGATACGCTCGGCGTGC 1740
QY	1741	AAACTCTCTCAAGGAGAGGTCAATTCAGATGCGGCTCAACCAATCTGATGTTGGG 1800
DB	1741	AAACTCTCTCAAGGAGAGGTCAATTCAGATGCGGCTCAACCAATCTGATGTTGGG 1800
QY	1801	TGCGAATCTCCATGCGAGCCGGAACCGGATGTAAGAGTGTCACTTCAATGCTCACCGAC 1860
DB	1801	TGCGAATCTCCATGCGAGCCGGAACCGGATGTAAGAGTGTCACTTCAATGCTCACCGAC 1860
QY	1861	CCCTCCCAATCAACAGAGAGCGGCTAAGCGAGGCTGGCCAGGGGGTCTCCCGCTCC 1920
DB	1861	CCCTCCCAATCAACAGAGAGCGGCTAAGCGAGGCTGGCCAGGGGGTCTCCCGCTCC 1920
QY	1921	TTGGCAGCTCTTCACTAGCCAGTTGTCTGCGCTTCTCGAAGCGCATATTAAC 1980
DB	1921	TTGGCAGCTCTTCACTAGCCAGTTGTCTGCGCTTCTCGAAGCGCATATTAAC 1980
QY	1981	CAAAATGACTTCCCAACGATGATGAGGCGCACTCTGTGGGGGCAATGAGATG 2040
DB	1981	CAAAATGACTTCCCAACGATGATGAGGCGCACTCTGTGGGGGCAATGAGATG 2040

QY 2041 GGGGGGACATTAACCGCGTGAGTCAGAGAAAGAGTAATCTTGACCTTTTGAC 2100  
 |||||  
 Db 2041 GGGGGGACATTAACCGCGTGAGTCAGAGAAAGAGTAATCTTGACCTTTTGAC 2100  
 QY 2101 CCGCTCCGAGCGGAGAGATGAGCGGAGAGTCCGTCGCGGAGATCTTGCGGAAA 2160  
 |||||  
 Db 2101 CCGCTCCGAGCGGAGAGATGAGCGGAGAGTCCGTCGCGGAGATCTTGCGGAAA 2160  
 QY 2161 TCCAGAAATTTCCACAGGATGCGCGGATGCGGATTAACACCTTCGCGT 2220  
 |||||  
 Db 2161 TCCAGAAATTTCCACAGGATGCGCGGATGCGGATTAACACCTTCGCGT 2220  
 QY 2221 CTGAGATCTTGAGAGGCGCGGACTAGTCCCTCAGATGATAGTGGTCCCACTGCA 2280  
 |||||  
 Db 2221 CTGAGATCTTGAGAGGCGCGGACTAGTCCCTCAGATGATAGTGGTCCCACTGCA 2280  
 QY 2281 CCTACTAAGACCCCTCTATACCACTTCAACGAGAGAGAGACAGTTGTTCTGACGAA 2340  
 |||||  
 Db 2281 CCTACTAAGACCCCTCTATACCACTTCAACGAGAGAGAGACAGTTGTTCTGACGAA 2340  
 QY 2341 TCCACCGTGTCTTGCGCGGAGCTTGCCACAAAGGCTTTCGTAAGTCCGAAACG 2400  
 |||||  
 Db 2341 TCCACCGTGTCTTGCGCGGAGCTTGCCACAAAGGCTTTCGTAAGTCCGAAACG 2400  
 QY 2401 TCGGCGGTGACAGCGGACCGCAACCGCCCTCTGACCAACCTTCCGACGAGCGGCA 2460  
 |||||  
 Db 2401 TCGGCGGTGACAGCGGACCGCAACCGCCCTCTGACCAACCTTCCGACGAGCGGCA 2460  
 QY 2461 GCAGATCTGACGTTGATGTAATCTCTCATGCCCCCTTGAGAGGAGCGGAGGAC 2520  
 |||||  
 Db 2461 GCAGATCTGACGTTGATGTAATCTCTCATGCCCCCTTGAGAGGAGCGGAGGAC 2520  
 QY 2521 CCCGATCTCAGCGACGGGTCTTGCTCAACCTGATGAGAGAGCGGATGAGACCTGTC 2580  
 |||||  
 Db 2521 CCCGATCTCAGCGACGGGTCTTGCTCAACCTGATGAGAGAGCGGATGAGACCTGTC 2580  
 QY 2581 TGTGCTCATGATCTCTACACATGAGACAGCGCTCTGATCAACGCAATGCGTGGAGGAA 2640  
 |||||  
 Db 2581 TGTGCTCATGATCTCTACACATGAGACAGCGCTCTGATCAACGCAATGCGTGGAGGAA 2640  
 QY 2641 AGCAAGCTGCCATCAACCGGTTGAGCAACTTTGCTGCGTCAACCAATGATGTCTAC 2700  
 |||||  
 Db 2641 AGCAAGCTGCCATCAACCGGTTGAGCAACTTTGCTGCGTCAACCAATGATGTCTAC 2700  
 QY 2701 GCTACACATCCCGACCGGACAGCGGACAGAGAAAGTCACTTTGACATGTCGAA 2760  
 |||||  
 Db 2701 GCTACACATCCCGACCGGACAGCGGACAGAGAAAGTCACTTTGACATGTCGAA 2760  
 QY 2761 ATCTTGAGAGATCACTACAGAGAGTGTCAAGAGAGAGAGGAGGCTCCACAGTT 2820  
 |||||  
 Db 2761 ATCTTGAGAGATCACTACAGAGAGTGTCAAGAGAGAGAGGAGGCTCCACAGTT 2820  
 QY 2821 AAGGCTAAGCTTCTATCAGTAGAGAGAGCTTGACGCGCCCACTTCGCGCAAA 2880  
 |||||  
 Db 2821 AAGGCTAAGCTTCTATCAGTAGAGAGAGCTTGACGCGCCCACTTCGCGCAAA 2880  
 QY 2881 TCTAAATTTGGCTATGAGGAGAGAGAGCTCCGAACTTACGCAAGGCGCATTAACAC 2940  
 |||||  
 Db 2881 TCTAAATTTGGCTATGAGGAGAGAGAGCTCCGAACTTACGCAAGGCGCATTAACAC 2940  
 QY 2941 ATCCGCTCGTGTGAGAGAGCTTGTGAGAGACACTGAACCAATTTGACACACATC 3000  
 |||||  
 Db 2941 ATCCGCTCGTGTGAGAGAGCTTGTGAGAGACACTGAACCAATTTGACACACATC 3000  
 QY 3001 ATGCAAAAATGAGGTTTTCTGCTCAACAGAGAGAGAGCGGCAAGCCAGCTCGC 3060  
 |||||  
 Db 3001 ATGCAAAAATGAGGTTTTCTGCTCAACAGAGAGAGAGCGGCAAGCCAGCTCGC 3060  
 QY 3061 CTTATGCTGTCCCAACTTTGGGGGCTCGGTGTGAGAGAGAGAGAGAGAGAGAGAG 3120  
 |||||  
 Db 3061 CTTATGCTGTCCCAACTTTGGGGGCTCGGTGTGAGAGAGAGAGAGAGAGAGAGAG 3120  
 QY 3121 GTCTCACCCCTCCCTCAGGCTGTGATGGGTCTCTCGTAGAGATTCAGATTTCCCGGA 3180  
 |||||

Db 3121 GTCTCACCCCTCCCTCAGGCTGTGATGGGTCTCTCGTAGAGATTCAGATTTCCCGGA 3180  
 |||||  
 QY 3181 CAGCGGTGAGATTTCTGTTGAAAGCGCTGAAATCAAGAAGACCTTATGGCTTTGCA 3240  
 |||||  
 Db 3181 CAGCGGTGAGATTTCTGTTGAAAGCGCTGAAATCAAGAAGACCTTATGGCTTTGCA 3240  
 QY 3241 TATGACACCGGCTTTTGAAGTCAAGTCACTGAGTAATGACATCCGTGTAGAGAGTCA 3300  
 |||||  
 Db 3241 TATGACACCGGCTTTTGAAGTCAAGTCACTGAGTAATGACATCCGTGTAGAGAGTCA 3300  
 QY 3301 ATTATCAATGTTGTGACTTGGCGCCGGAAGCCAGACAGGCTAAGATGCTCAGAGAG 3360  
 |||||  
 Db 3301 ATTATCAATGTTGTGACTTGGCGCCGGAAGCCAGACAGGCTAAGATGCTCAGAGAG 3360  
 QY 3361 CGGCTTATATCGGGGCTCCCTGACTAATTCAAAGGAGAGATGCGGCTATCGCCG 3420  
 |||||  
 Db 3361 CGGCTTATATCGGGGCTCCCTGACTAATTCAAAGGAGAGATGCGGCTATCGCCG 3420  
 QY 3421 TGGCGGCGAGGGGCTGAGAGTGAAGTGGATGAGTGAATACCTTCAATGTTACTTGAG 3480  
 |||||  
 Db 3421 TGGCGGCGAGGGGCTGAGAGTGAAGTGGATGAGTGAATACCTTCAATGTTACTTGAG 3480  
 QY 3481 GCCTTGACGCTGTGAGAGCTGCAAGCTCAGAGCTGACATGCTGTGCGGAGAG 3540  
 |||||  
 Db 3481 GCCTTGACGCTGTGAGAGCTGCAAGCTCAGAGCTGACATGCTGTGCGGAGAG 3540  
 QY 3541 GGCCTTGCTTATCTGTGAGAGCGGGAACCCAGAGAGAGCGGCGAGCTTACGAGTC 3600  
 |||||  
 Db 3541 GGCCTTGCTTATCTGTGAGAGCGGGAACCCAGAGAGAGCGGCGAGCTTACGAGTC 3600  
 QY 3601 TTCAGGAGAGTATGACTGATGATGCTGCCCCCGGGGAGCCCGGCAACAGAAATAC 3660  
 |||||  
 Db 3601 TTCAGGAGAGTATGACTGATGATGCTGCCCCCGGGGAGCCCGGCAACAGAAATAC 3660  
 QY 3661 GACCTGAGATGATATCATCATGCTCTCAATGTTGCTGCGGAGAGATGATGATGAC 3720  
 |||||  
 Db 3661 GACCTGAGATGATATCATCATGCTCTCAATGTTGCTGCGGAGAGATGATGATGAC 3720  
 QY 3721 AAAAGGATATCTACTCAACCGGTGACCG 3750  
 |||||  
 Db 3721 AAAAGGATATCTACTCAACCGGTGACCG 3750

RESULT 4  
 AFJ13916 9359 bp mRNA linear VRL 01-JAN-2002  
 LOCUS  
 DEFINITION Hepatitis C virus polypeptide mRNA, partial cds.  
 ACCESSION AFJ13916  
 VERSION AFJ13916.1 GI:118027684  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Hepatitis C virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.  
 REFERENCE  
 1 (bases 1 to 9359)  
 Fanning, L.J., Itakura, J., Nagayama, K. and Enomoto, N.  
 Characteristics of Hepatitis C viral genome associated with disease  
 progression in a homogeneous patient population  
 Unpublished  
 JOURNAL  
 REFERENCE  
 2 (bases 1 to 9359)  
 Fanning, L.J., Itakura, J., Nagayama, K. and Enomoto, N.  
 Direct Submission  
 Submitted (17-OCT-2000) Medicine, National University of Ireland,  
 Cork, Hepatitis C Unit, Clinical Sciences Building, Cork University  
 Hospital, Cork, Ireland  
 FEATURES  
 Location/Qualifiers  
 1..9359  
 /organism="Hepatitis C virus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:11103"  
 /note="Isolated from source plasma identified as the  
 single source of the iatrogenic infection of Rheus

5'UTR  
CDS

negative women in Ireland from May 1977-November 1978;  
exposure was through HCV 1b contaminated anti-D  
immunoglobulin"

1. .329

330. .&gt;9359

/codon\_start=1

/product="polyprotein"

/protein\_id="AL55821.1"

/db\_xref="GI:18027685"

/translation="MSTNPKPKQKTKGNTNRKPPQVKKPFGGQIVGGVYLIPRRGRRL

GVATRKTSERSQPRGRKQPIPAKRPBEGTMAQPGIPMLTIBENMGWMLISRG  
SRPSWGPDPBRRSRNLGKVIDLTCGFADLMGIVPLGAPGLGALAHAGRVLEAD  
GVNATGNLPGCSFSIFLLALISCLTIPASALEVRNVSQVHTVNDCSNATYVEAD  
MIMHTGCVCEVENNSRRCVATPLTAARNSIPTTTIRRDVLDMMMLSPFACSMY  
VGLDGSVFLVSLQPTFSPRHEVTDKNCISIGHNSGRHMAVMMMSMTTALVY  
SOLRIPOAVVMVAGAHWCYAGLAAYSMAGWMAVYLIMLIPAVDEPTTGGAO  
AHTTGLATLTFSPESQNTQILINTGSMHNTALNCNDSLNTGFLAALFYARFRAS  
GCEBRNASCRPIDKFAQGWGPIAAKPLSIDQPCYCMHAPQCGIVPASQVCGPYAC  
FPSPVVGTTRDFGVPTYSWGENEDVLLNTRPRRGMFGCTMMNGTGFKTGCG  
PCPDIGVGNLTLICTDCEPKRKEATYTCGSPMLTRCMWDYFURLMHPCTVNF  
TIKVMYVGVGHRHLAACNTRGRCDLDRSELSPILSTTEMOVLPSCSPFTL  
PALSTGLIHHRIVDOYLYXGSAVSAIKREYVLLPLLAARPCACIMMLL  
IYQAEALENVVLMAASVAGHTGIIISFLVFCAMYTIKRILVGAAYFYGWPIIL  
LIALPRAVAMDRMAASCGAVFTGLALTLSPHYKVFALAILMWLOFYTRAEAH  
LQWIPPLNVRGRDAIILMCAIHPELIPATIKIILAGLMLVLAQAGITRPFYER  
AHGLIACVLRVAGGHVYQALMLALITGVVDHLTPLDMHEGLRLDAVVE  
PVSFDMETKILITWGA DTAA CGDIILGLPSARGREILGPDLSLEGQWRLIADIT  
AYAQOTRGLGCTVSLTGRDKNVRGEVQVNSTATQSPALTCVNVCMVTHGAAATK  
TLAQGKPIITQMTNVQDLVGLMAPGASLTPCTGSDLLVLRHADVLPVRERG  
DSRGSLSLSPRVSYLKGSSGGLCPSGNAVIGFRAVCAVGAAYLVLPVPSMRT  
MRSPVFTDSSPAPVQTFQVAHLHAPTGSGKSTKVAIPAAGVAVLVPVPSMRT  
GFGAYSKAHGVDPNRTGVRTITGGAITYSTYKFLADGGSGGAGVILIDEGHS  
TOSTTIGITGVLDQAEAGALVLAATAPGSAVTVPHPNIEVALNSNGETPFYKX  
AIIETIKGRHLIFCHSRKCKDELAKLSGLMANVAYIRGDAVITPSSGVVYVA  
TDLMTGFTGDFDSVIDCNTCVIQTDFSLDPTFTIETTVPODVAISRSORGTIRG  
RTGIYRFVTPGERSGMFDSSVLCCEYDACAMAYELTPASTRLAYLNTPELPCQ  
DHLFEWEGVFTGLIHDAFLISQTKQADGFPYLVAQAVCAAPPSMMOMK  
LILKPLTHGPFTPLRYLGAONEVILITHEITKEIIMACMDALEVATSTVILVGYLA  
ALAAVCLTGGSVIVIGRIILSGKPAIVDEEVLXYODEMECEATLPIYEOGMOLAE  
QFKOKALGLIQTAKQAEAAPIVESKWRALBPKAKMNPISGLOIYALSLTLN  
PATASLMFTASTISPLITTOHTLIFNLGSMVAAQLPSPASAFISGAPLAGAUSI  
GLAKVLDVILAGGAGVAGALVAFKVMGSPSTEDVNLPLIFISGALVGVVCA  
ILRRHVPEGGAVOMNMRILAFASRGHNVSPTHVPSDAAALITLSLITITOLK  
RLHOMINEDCSTPCSSMLDVMDCITVLTDFKMLQSDVLAIRLPGVPLSCORXYK  
GVMRGDMGIMOTPCGACITGVHNGSMRIVGRTCSNTHGTFPIINATYCTGSP  
ANYSRALMPVTAEBYLAEYQVGDPHYVNMPTDNTKPCQVAPAPREFVDPVRLHR  
YAPCKPLREEVTFQVGLNQYLVGSLPCEPBDVAVLTSMLIDSHITAEAKRL  
ARSPSPILASSASQLSAPSLKATCTTRHSDPADLLEANLMDRQEGNITVBSN  
KVVILDSFELRAEBDEBEREVSVAELIRSRKPRAMPPIWARDYNPPLVESKBDY  
VPPVHGCPLPTKAPLIPPRKRTVVLTESTYSSALAEATKTGSSSSSDVCSGT  
ATA PPDPSDDGAGSDVESMSMPLEGRDPLSDGSMSTVSEASDDVVCMS  
YTMGALITPCSAEBSKLPINALSNLSLRHNNVYATSSASQOROKTTFEDLOQLD  
DHYRDVLKEMKAKASTVKALPLVEBECKLPSPASKSKFGYKADNRINLSAVNHI  
LSVWKDLIDETETPIDTTIMAKNEVFCVQBEKGRPARLIVPPDGLVRCERMAIYD  
VNSTLPQAVMSSSYGQYSPQORVEPLVPLNKKKKKMGAVYTRFDSVSTDIRV  
EBSIYOCCLAPBAROARISLRETLVGGELTNSKQNCYCRKSRASGVLTSCGTL  
TVYKLSAACRAKLDCTMLVCGDDLVVICSAGTEDASLRVFEAMRTASAPG  
DPKPEYDELELITSCSNVSAVDHASKRYIYITRPTTLAANAETHAVTNVSM  
GNITMTPITLMARMLMTHTFSTILAOBOLKALDOCTYAGCSIEPLDLPQITQRH  
GLSAFSLHASYSGEINRVACLRLGAPLRVHRHRSVRALILQGGGAACGYL  
FNNAVRTKLKLTPIPAASQLDLSWFAVAGSGDIYHLSRARPRFMLCLILSVGV  
GTYLLEPR"

ORIGIN

Query Match 88.5%; Score 3319.8; DB 14; Length 9359;

Best Local Similarity 92.8%; Pred. No. 0;

Matches 3480; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

QY 1 TGGAGGGCGCTCTTCAAGGCTTCAAGGCTGAGTATCCCACTTCGTCCTCCAAACAAG 60  
|||||  
DB 5001 TGGAGGGCGCTCTTCAAGGCTTCAAGGCTGAGTATCCCACTTCGTCCTCCAAACAAG 5060  
|||||

DB 5061 CAGGACGAGGAGACAATCTCCCTTAATTGGTAGATACAGGCTACGGTGGCCAGGGCT 5120  
|||||  
QY 121 CAGGCCCCCACTTCATCATGGGATGATAAATGTGGAAGTGTCTATGCGGTTAAAGCTACT 180  
|||||  
DB 5121 CAGGCTCCACCCCATCGTGGACCAAAATGTGGAATGTCTATGCGGTTAAAGCCACG 5180  
|||||  
QY 181 CTGCGGGGGCCAAACCCCTTGTGTATAGCTGGAGCCGTCMAAACGAGGTCAACCTTC 240  
|||||  
DB 5181 CTGCAAGGGGCCAACCCCTGTGTATAGGCTAGGAGCGGTCMAAATGAGGTCAATCTTC 5240  
|||||  
QY 241 ACAACCCCATTAACAATTCATCATGGCATGATGACGCCGACTGGAGGTCTGCACG 300  
|||||  
DB 5241 ACAACCCCATTAATAATTCAATCATGGCATGATGACGCTGATCTGGAGGTCTGCACG 5300  
|||||  
QY 301 AGCAGCTGGGTCGTGGTGGGGGGGCGTTCGAGCTCTGAGCTGCGATATGTTGACAACA 360  
|||||  
DB 5301 AGCAGCTGGGTCGTGTAGGCGGAGTCTTCGAGCTCTGGCGGATATGCTGACACACA 5360  
|||||  
QY 361 GCGAGCGTGTGATTTGTGGGTAGATCATCTTGTCCGGCGGCGGCTATTTGTTCCGAC 420  
|||||  
DB 5361 GCGAGCGTGTGATTTGTAGGAGATCATCTTGTGGAAGCGGCGCTATTTGATCCGAC 5420  
|||||  
QY 421 AGGGAATCTCTTACAGGAGTTCATGATGATGGAAGATGGCGCTGCACTCCCTTAC 480  
|||||  
DB 5421 AGGGAATCTCTTACAGGAGTTCATGATGATGGAAGATGGCGCACACACTCCCTTAC 5480  
|||||  
QY 481 ATCGAGCAGGGAATGAGTCTCCCGGAGCTTCAAGCAAAAAGCGCTGGGTTGTCGAG 540  
|||||  
DB 5481 ATCGAACAAAGGAATGAGTCTCCCGGAACATTCAGCAAAAGCAATCGGGTTGTCGAG 5540  
|||||  
QY 541 ACAGCACCACAAAGCGGAGCGCGTCTCCCGTGTGTGAGTTCAAAGTGGCGAGCCCTT 600  
|||||  
DB 5541 ACAGCACCACAAAGCAGAGCGGTGCTCCCGTGTGTGAGTTCAAAGTGGCGAGCCCTT 5600  
|||||  
QY 601 GAGACTTGTGGGCGAAACAATGTGGAATTCTATCAGCGGGAATCAATATCTTAAGCAGGC 660  
|||||  
DB 5601 GAGAGCTTGTGGGCGAAACAATGTGGAATTCTATCAGCGGGAATCAATATCTTAAGCAGGC 5660  
|||||  
QY 661 TTGTCCACTCTGCTGGGGAATCCCGGATGTGATCATGATGAGCGGTTCAAGCCTGTGC 720  
|||||  
DB 5661 TTGTCCACTCTGCTGGGGAATCCCGGATGTGATCATGATGAGCAATTCACAGCCTTATTC 5720  
|||||  
QY 721 ACTAGCCGCTCAACCAACCTTCACTCTGCTTAAATCATCTGCGGGGATGGATGACC 780  
|||||  
DB 5721 ACAGGCGGCTTACACCAACCTTCACTCTGCTTAAATCATCTGCGGGGATGGATGACC 5780  
|||||  
QY 781 GCCCAATCTGCTCCCCAGTGTGCTTCACTTGTGATGAGCGCGGCAATGCTGTGTCG 840  
|||||  
DB 5781 GCCCAATCTGCCCCCAGCGCTGCTTGTGATGAGCGCGGCAATGCTGTGTCGCGCG 5840  
|||||  
QY 841 GCTGTGGGAGCATAGGCTTGGGAAGTGTGGAATCATGCTGGCGGGCTATGAGAGA 900  
|||||  
DB 5841 GCGGTGTGAGCATAGGCTTGGGAAGTGTGGAATCATGCTGGCGGGCTATGAGAGA 5900  
|||||  
QY 901 GGAAGTGGAGGCGCTGCTGAGCTTAAAGTATAGAGCGGCGAAATGCTTCACCGAG 960  
|||||  
DB 5901 GGGGTGGAGGCGCTCTGAGCTTAAAGTATAGAGTGGGAGATGCTTCACCTGAG 5960  
|||||  
QY 961 GACCTGTTAATTAATCTCCCTGCACTCTCTCTCTGTGTCCTGTGCTCGGCGTCTG 1020  
|||||  
DB 5961 GACCTGTTAATTAATCTCCCTGCACTCTCTCTCTGTGTCCTGTGCTCGGCGTCTG 6020  
|||||  
QY 1021 TGGCAGCAGATCTGCTGGGAGCGTGGGATCCAGGGGAGGGGGCTGTGAGTGGATGAAC 1080  
|||||  
DB 6021 TGGCAGCAGATTAATGCTGGGATGTGGGCCCAAGGGGAGGGGCTGTGAGTGGATGAAC 6080  
|||||  
QY 1081 CGGCTATATAGCTTCCCTCGCGGGGTAACCAATGTTTCCCGCAGCACTATGTGCCAGAG 1140  
|||||  
DB 6081 CGGCTATATAGCTTCCCTCGCGGGGTAACCAATGTTTCCCGCAGCACTATGTGCCAGAG 6140  
|||||  
QY 1141 AGGAGCGCGCAGACGTTGCACTGATCTCTCGACCTTAATATCAACCACTGTTG 1200  
|||||



Db 6141 AGCAGCGCTGACGCGCTGCTCACTCAGATCTCTCCAGCCTTACCATCACTCAGCTCTG 6200  
Qy 1201 AAGAGGCTCCACCGAGGATTTAAAGAGAGAGCTGCTCCAGCCCTGCTCGGCTGCTGCTA 1260  
Db 6201 AAGAGGCTCCACCGAGGATTTAAAGAGAGAGCTGCTCCAGCCCTGCTCGGCTGCTGCTA 6260  
Qy 1261 AAGGATGTTTGGAGCTGATATGACAGTTTGGCTGACTTCAAGACCTGGCTCAGATCC 1320  
Db 6261 AAGGAGCTTTGGAGCTGATATGACAGTTTGGCTGACTTCAAGACCTGGCTCAGATCC 6320  
Qy 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCAACGCTGGATCAAGAGG 1380  
Db 6321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCAACGCTGGATCAAGAGG 6380  
Qy 1381 GTCTGGGGGGGAGACCGGATCATGAGCAACCTGCTCATGTGAGAGCAAGATACCGGA 1440  
Db 6381 GTCTGGGGGGGAGACCGGATCATGAGCAACCTGCTCATGTGAGAGCAAGATACCGGA 6440  
Qy 1441 CATGTCAAAAACGTTTCCATGAGGATGTTGGGCTTAAGACCTGTAGTAAATATGTGCA 1500  
Db 6441 CATGTCAAAAACGTTTCCATGAGGATGTTGGGCTTAAGACCTGTAGTAAATATGTGCA 6500  
Qy 1501 GGAACATTTCCCATCAACGATACCAACGAGGCTCTGACGCTTCCCAAGCGCAAC 1560  
Db 6501 GGAACATTTCCCATCAACGATACCAACGAGGCTCTGACGCTTCCCAAGCGCAAC 6560  
Qy 1561 TATTCAGGGGCGCTGAGCGGCTGCTGCTGAGAGAGTGAAGAGTTAAGCGGCTGGG 1620  
Db 6561 TATTCAGGGGCGCTGAGCGGCTGCTGCTGAGAGAGTGAAGAGTTAAGCGGCTGGG 6620  
Qy 1621 GATTTCCACTACGTGACGAGCATGACCATGACAACTAAATATGCGCTGACAGTTTCA 1680  
Db 6621 GATTTCCACTACGTGACGAGCATGACCATGACAACTAAATATGCGCTGACAGTTTCC 6680  
Qy 1681 GCGCCGGAATTTCTTCAAGAGTGAAGTGGGCTGCAAGTCAAGTCAAGTCAAGTCAAG 1740  
Db 6681 GCGCCGGAATTTCTTCAAGAGTGAAGTGGGCTGCAAGTCAAGTCAAGTCAAGTCAAG 6740  
Qy 1741 AAACCTCTCTTACGAGGAGAGTCACTTCAAGTTCGAGGCTCAACCAATACCTGCTGG 1800  
Db 6741 AAACCTCTCTTACGAGGAGAGTCACTTCAAGTTCGAGGCTCAACCAATACCTGCTGG 6800  
Qy 1801 TCCAGACTCCCATGAGCGGCAACCGAGTGTAGACAGTGTCACTTCCATGCTCAACGAC 1860  
Db 6801 TCCAGACTCCCATGAGCGGCAACCGAGTGTAGACAGTGTCACTTCCATGCTCAACGAC 6860  
Qy 1861 CCTTCCCATCAACGAGAGAGCGCTAAGCGAGCTGAGCGAGGCTGCTCCCTCC 1920  
Db 6861 CCTTCCCATCAACGAGAGAGCGCTAAGCGAGCTGAGCGAGGCTGCTCCCTCC 6920  
Qy 1921 TTGGCCAGCTCTTACGCTAGCAGTGTCTGCGCTTCTTCAAGGCGGACATACATTAAC 1980  
Db 6921 TTGGCCAGCTCTTACGCTAGCAGTGTCTGCGCTTCTTCAAGGCGGACATACATTAAC 6980  
Qy 1981 CAATAATGACTTCCAGAGCTGACCTCATGAGGCAACCTCCGTTGGGCGGCAATGAGAT 2040  
Db 6981 CAATAATGACTTCCAGAGCTGACCTCATGAGGCAACCTCCGTTGGGCGGCAATGAGAT 7040  
Qy 2041 GGGGGGAGCATTAACCGCGTGAAGTCAAGAGCAAGGTAGTAACTCTTGAC 2100  
Db 7041 GGGGGGAGCATTAACCGCGTGAAGTCAAGAGCAAGGTAGTAACTCTTGAC 7100  
Qy 2101 CCGCTCCGAGCGGAGAGTGAAGCGGAAAGTGTCCGTCCCGGCGAGATCTTGGGAA 2160  
Db 7101 CCGCTCCGAGCGGAGAGTGAAGCGGAAAGTGTCCGTCCCGGCGAGATCTTGGGAA 7160  
Qy 2161 TCCAGAAATTTCCACAGGATGCGGATGAGGAGCGCGGATTTCAACCTCTGCTG 2220  
Db 7161 TCCAGAAATTTCCACAGGATGCGGATGAGGAGCGCGGATTTCAACCTCTGCTG 7220  
Qy 2221 CTGAGATCTGAGAGCGCGGAGCTAGTCTCTTCAATGTGTACATGTGGTCCCATGCA 2280  
Db 7221 CTGAGATCTGAGAGCGCGGAGCTAGTCTCTTCAATGTGTACATGTGGTCCCATGCA 7280

Qy 2281 CTTAATTAAGACCCCTCTTATACCACTTCAACGAGAGAGAGCAGTTGTTTGAAGAA 2340  
Db 7281 CTTAATTAAGACCCCTCTTATACCACTTCAACGAGAGAGAGCAGTTGTTTGAAGAA 7340  
Qy 2341 TCCAGCGTGTCTTCTGCGCTGAGCGGAGCTTGGCAAAAGCTTTCGTAGCTCGGAACG 2400  
Db 7341 TCCAGCGTGTCTTCTGCGCTGAGCGGAGCTTGGCAAAAGCTTTCGTAGCTCGGAACG 7400  
Qy 2401 TCGGCGGTGAGACAGCGGAGCGGACACCGCTCTCTGACCAACCTTCCAGACGCTGGA 2460  
Db 7401 TCGGCGGTGAGACAGCGGAGCGGAGCGGCTCTCTGACCAACCTTCCAGACGCTGGA 7460  
Qy 2461 GCAGAGTCTGAGTGTGATGTTCTCTTCAATGCCCCCTTGAAGGGGAGCGCGGAGGAC 2520  
Db 7461 GCAGAGTCTGAGTGTGATGTTCTCTTCAATGCCCCCTTGAAGGGGAGCGCGGAGGAC 7520  
Qy 2521 CCGATCTCAAGCAGCGGCTTCTGCTTACCGTGAAGTGAAGAGCGGATGAGACGTCGTC 2580  
Db 7521 CCGATCTCAAGCAGCGGCTTCTGCTTACCGTGAAGTGAAGAGCGGATGAGACGTCGTC 7580  
Qy 2581 TGTCTGCTGATGTTCTTACATGAGAGAGCGGCTTGTATCACGCTATGCTCGGAGAA 2640  
Db 7581 TGTCTGCTGATGTTCTTACATGAGAGAGCGGCTTGTATCACGCTATGCTCGGAGAA 7640  
Qy 2641 AGCAGGTGCGCATCAACGCGTTGAGCAACTCTTGTGCTGACACAAATGTCTAC 2700  
Db 7641 AGCAGGTGCGCATCAACGCGTTGAGCAACTCTTGTGCTGACACAAATGTCTAC 7700  
Qy 2701 GCTACACATCCCGCAGCGGAGAGCGGAGCGGAGAGAGTCACTTTGACAGATGCA 2760  
Db 7701 GCTACACATCCCGCAGCGGAGAGCGGAGCGGAGAGAGTCACTTTGACAGATGCA 7760  
Qy 2761 ATCTGAGAGATCACTACAGAGAGCTGCTCAAGAGATGAAGGCGAAGGCTTCAAGTT 2820  
Db 7761 ATCTGAGAGATCACTACAGAGAGCTGCTCAAGAGATGAAGGCGAAGGCTTCAAGTT 7820  
Qy 2821 AAGGCTTAAGCTTATACAGTGAAGAGAGCTGCAAGCTGACCGCCCACTTGGGCA 2880  
Db 7821 AAGGCTTAAGCTTATACAGTGAAGAGAGCTGCAAGCTGACCGCCCACTTGGGCA 7880  
Qy 2881 TCTAATTTGGCTATGAGGAGAGAGAGCTGCGGACCTTATCCAGAGGCTATTAACAC 2940  
Db 7881 TCTAATTTGGCTATGAGGAGAGAGAGCTGCGGACCTTATCCAGAGGCTATTAACAC 7940  
Qy 2941 ATCTGCTCGGTGAGAGAGCTTGTGGAAGACATGAAACCAATTTGACACACCATC 3000  
Db 7941 ATCTGCTCGGTGAGAGAGCTTGTGGAAGACATGAAACCAATTTGACACACCATC 8000  
Qy 3001 ATGGCAAAAAATGAGGTTTTTTCGCTCCCAACCAAGAGAGAGCGGCAAGCCAGCTCCG 3060  
Db 8001 ATGGCAAAAAATGAGGTTTTTTCGCTCCCAACCAAGAGAGAGCGGCAAGCCAGCTCCG 8060  
Qy 3061 CTTATGCTGTCCAGACTTGGGAGGCTCGGTGTGAGAGAAAAATGAGCCCTTATGAGCTG 3120  
Db 8061 CTTATGCTGTCCAGACTTGGGAGGCTCGGTGTGAGAGAAAAATGAGCCCTTATGAGCTG 8120  
Qy 3121 GTCTCAACCTTCCAGCTGATGAGGCTCTCTGATGAGATTCAGATTTCTCTGGA 3180  
Db 8121 GTCTCAACCTTCCAGCTGATGAGGCTCTCTGATGAGATTCAGATTTCTCTGGA 8180  
Qy 3181 CAGCGGCTGAGATTTCTGCTGAGAGAGCTGAGAGAAATCAAAAGAGCCCTTATGAGCTT 3240  
Db 8181 CAGCGGCTGAGATTTCTGCTGAGAGAGCTGAGAGAAATCAAAAGAGCCCTTATGAGCTT 8240  
Qy 3241 TATGACACCGCTGTTTGAAGTCAAGTCACTGAGAGATGACATCCGTGTAGAGAGTCA 3300  
Db 8241 TATGACACCGCTGTTTGAAGTCAAGTCACTGAGAGATGACATCCGTGTAGAGAGTCA 8300  
Qy 3301 ATTATCAATGTTGTGATGCTTGGCCCGAGAGCGAGCGCAATGAGTGTCTCACAGAG 3360  
Db 8301 ATTATCAATGTTGTGATGCTTGGCCCGAGAGCGAGCGCAATGAGTGTGTCTCACAGAG 8360

Oy		3361	CGGCTTTATTCGGGGGCCCCGTACATAATTCAAAAGGCGAAGACTGGGTATGCCCGG	3420
Dd		8361	CGGCTTATTCATCGGGGTCGCCCTGACTTAACAAGGGGCAGAATCGGGTTATCCCGG	8420
Oy		3421	TGCCGCGGACGGGGTGCTGTACGACATAGCTGCGGTAATACCCTCACATGTTACTTGAAG	3480
Dd		8421	TGCCGCGGACGGGGTGCTGTACGACACAGCTGCGGTAATACCCTCACATGTTACTTGAAG	8480
Oy		3481	GCTCTGCAGCCTGTGAGCTTGCAAAAGCTCCAGAAGCTGCAGATGCTGTGTGCGGAGAC	3540
Dd		8481	GCTCTGCAGCCTGTGAGCTTGCAAGCTCCAGAAGCTGCAGATGCTGTGTGCGGAGAC	8540
Oy		3541	GGCCTTGTCGTTATCTGTGAGAGCGCGGGAAAACCCAGAGAGAACGGCGGACGCTTACAGTC	3600
Dd		8541	GACCTCGTCGTTATCTGTGAGAACCGCGGAAAACCCAMGAGAACGGCGGACGCTTACAGTC	8600
Oy		3601	TTACAGGAGGCTATGACTAGTAGTCTGTGCCCCCGGGGAGACCGCGCCCAACAGAAATAC	3660
Dd		8601	TTACAGGAGGCTATGACTAGTAGTCTGTGCCCCCGGGGAGACCGCGCCCAACAGAAATAC	8660
Oy		3661	GACCTGAGTGTATACATCATGCTCTCCCAATGTGTGGTGGCGACGATGATCTGGC	3720
Dd		8661	GACCTGAGTGTATACATCATGCTCTCCCAATGTGTGGTGGCGACGATGATCTGGC	8720
Oy		3721	AAAGGGTATCTACTACATCCACCCGTCAGCC	3749
Dd		8721	AAAGGGTGTACTACTACATCCCGTGATCCC	8749
RESULT 5				
LOCUS	HPCK1S2			
DEFINITION	Hepatitis C virus (strain HCV-1b, clone HCV-KI-S2), complete genome sequence.	9410 bp	RNA linear	VRL 10-FEB-1999
ACCESSION	D50485			
VERSION	D50485.1 GI:1030704			
KEYWORDS	polyprotein; core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A; NS5B; envelope protein; non-structural protein; interferon-sensitive; interferon-resistant; IFN-sensitive; IFN-resistant; ISDR; interferon sensitivity determining region; HVR; hypervariable region.			
SOURCE	Hepatitis C virus			
ORGANISM	Hepatitis C virus			
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			
REFERENCE	Enomoto,N., Sakuma,I., Asahina,Y., Kurosaki,M., Murakami,T., Yamamoto,C., Izumi,N., Marumo,F. and Sato,C., Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region J. Clin. Invest. 96 (1), 224-230 (1995)			
TITLE				
JOURNAL				
MEDLINE	95340824			
PUBMED	7542279			
REFERENCE	2 (bases 1 to 9410)			
AUTHORS	Enomoto,N.			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 9410)			
AUTHORS	Enomoto,N.			
TITLE	Direct Submision			
JOURNAL	Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental University, Second Department of Internal Medicine; 1-5-45 Yushima, Bunkyo-ku, Tokyo 113, Japan (E-mail:PXN04522@niftyserve.or.jp, Tel:03-3813-6111(Ex.3224), Fax:03-3818-7177)			
FEATURES	Location/Qualifiers			
SOURCE	1..9410			
	/organism="Hepatitis C virus"			
	/mol_type="genomic RNA"			
	/strain="HCV-1b"			
	/db_xref="taxon:11103"			
	/clone="HCV-KI-S2"			
	/note="IFN-sensitive HCV from patient 2"			
CDS	330..9362			

/codon\_start=1  
 /product="pol1yprotein"  
 /protein\_id="BA09076.1"  
 /db\_xref="gi:1814089"  
 /translation="MSTNTPKPKORTKNTNRRPDVYKPGGGQIVGGVYLLIPRRGPRLL  
 GVRATKTSSESGRRGRORPPIKARQPEGMAQMPYPMPLYGNEGGMAGMILLPRG  
 SRPMSPGDPERRSRNIGKVIDTLTGCFADLMQIPLVAGAPLGASALALAHGVLESD  
 GVNATATGNLPGCSFSIPLALLSLCTLPASAEVRNAGSYVHTNQCNSATSYEAD  
 MIHTPGCVPCPEANSSRCVATLPTPLANSSVPTTTIRHYDLLVGAACASAMV  
 VGDLCGSFVLVSQLPFSPRHETVQDCNSIPGHVSGHRMAMDMNNSPTTALV  
 SOLDIRPAVDNDVAGAHMVLAGLAYSVMGMKAVLIMWLPAGVGGATTTVGAQO  
 GHTTORLASETTPGAQORIOLVNTNGSMINRTALNCNDLSLNTGLALLETTRFMS  
 GCPERRIASCRSIDAPAGOMGPIYAEPSGSDORPCYHVARPCGQIPAEVCGSPYCS  
 FTSPSYVGGTIDRSAGPTYSNGENETDYLLNNTNAPAGCMWCPMMNGCPFKTGG  
 PPCNIGVGNNTLTCTPDCERKPEATYIKGSGSPMLTPRCMDYPRLMHYCTYAF  
 TIFKRVYMGVEHRNLNAACMWTEGRCDELDERRSELSPLLSTTEMQLPCCSPFTL  
 PVLSTGLIHQNIIVDOYUYGVGSAVVSFVIMKEVYLLPLFLIADARVCACLMMILL  
 IAOAALAEVNTVMNAAVAGAHGFLFLVFPCFCAVIRKGLVGAAYEAYGWPILL  
 LLIALPRAVYMDREMAASCGGAVFVALITLSPHKVLAHLIMWLQVYIRAEAC  
 LQWIPRLNATNGRDAIILLCAHPELISITKIYVAILGIPMTVMAQIGITKMPYPR  
 AQLGRACMLVAGVGHYVOMAFKLAALGVYDHLPLDQMAAGRLDRLAVAE  
 PVPSPDEMTKLIITWGAADIACGDIILGLAASARGKAIPLGAPDSLEGSMRLALPT  
 AYSQGTRELICCIITSLGRDNKNGVEVQVSAVATSPGLATCNGCMVYVAGSG  
 TLASGKLTITOMYTNVODLVGMOAPSGASLPTCGSSDILYLTSHAVIYVRRSG  
 DGRGLSPRVYSYKSGSSGGPLCPSGHAGVIRAVCRGKAVCAVDPIVMSMETL  
 MRSPIFTDNSSPRVAPVOTFOYAHLAHPTSGSKSKVPEVYEAAGCYKVLVYNSVAAITL  
 GPGAAMSFAHVDVPIRTGVRTITTGAPITYSTYTKGLADGSGGAYDIIIDECBS  
 TDTSTIIdGIITVLQAEFAGARLVLAATAPPGSVYVPHNIEEVALSNIGETIPYK  
 AIPLETTIKGSHLIFCHSKKKCDLAAKLSGLINAAVYRGDVSVIPASGVVVA  
 TDLATGTPGFDSDVIDCNTCVTQDVPFSLDPTPIETTVYPODVAERSQRGRGTGG  
 RAGIYRPTGPBERPSGMDSSVLCGYDAGCAWELPAEFSVRLRAYLNTPGIPVMO  
 DHEPMESEVPFGUTHIHAPLISOTKQAGNDPRLYVNOAYQVCAQAQAPPSMOMKX  
 LIRLKEPLTGLPRLYRIGAVQNVILHPIPTKIYIMGADLEVTSTVLYGVCLA  
 ALAACTLTGVSIVYVGRILISGKPAIIPDRVLIYREDEMEECSHLPIYEQMQLAE  
 OFKQALADLTOTAKOEAAPAVVESKQALETFWAGMNNVTSGLIYGLSTLPEN  
 PAISAMIAFTASISPLTTOHTLLEFNILGEGVAAQMPRAASAFVAGIAGAAVGSII  
 GLKXVLDLIGAGVAGVAGVAFKMGSMSTREDVLLPILISGALVVGCA  
 ILRRHVGBEAGVQMMARLIPASRGHNSPHTYVLSLAARPTQLSSLTITQLLK  
 RLHOMINDSCSTPCSGSWLRDYMVICLVLDTFTYLOSRLPRLPEVPLSCORGT  
 GVMRWDGIMOTTCTPCGAQITGHNXGSMRITVGPRTCSNTMHWTFPIAAYTTGCTSP  
 APNYSBALMRVAAEYEVETRVGDPHYVTGTTDNLKPCQVAPBEFKLDDGRRLH  
 YAPASKTLRBEVTFQVGLNOYVGSOLPCPEBEDVAVLSMLTDPSHITAEAKRLL  
 ARGSPPLASSASQASQASPKATCTTRHSPDAILENLIMROEGMNTIYESNS  
 KVVILDSERPLAEDEKEVSLPAEITPKSKKPPMPINARPDYNPILBSKDDPY  
 VEPVHGGCPLEPTKAPRIPERRRATVLETSVSSALAEIATVTESSSSAADSTY  
 ATAPVHGGCPLEPTKAPRIPERRRATVLETSVSSALAEIATVTESSSSAADSTY  
 YTMPTALITPCAAEBSKLPINALNSNLERHNNYVATTSASQROCKVTFDRLOVLD  
 DHYRVLKEMKAKASYKARLISVEBAKCLIPHSASKQYQACQDVRNLSQAINI  
 RSVWDILDEPTEPDIITTIKAKNGEPCVQPEKGRKAPRLIVPRDGLVRCCEKVALY  
 VVSTIPLQAVMPSTFGFQSPQQRVBYFLNMAKSKKCMGRVATPRDCDSTVTSQDL  
 BESTIYPLQAVMPSTFGFQSPQQRVBYFLNMAKSKKCMGRVATPRDCDSTVTSQDL  
 TCYUATVATCAAKAKLQDCITMLVCGDDIVLYCESAGTEBDASLVAFTEAMRTVASPG  
 DPQEPYULEITSCSSNVSAHDSAGVLYYLRDPDPLARMAHATHTFVNSVL  
 GSIIMAYETLWARMIMTHMPSILIAQOPELADCOIYGCVSIEBDDPOLIQRHL  
 GIASISLHYSPEGEINRVAASCLRLRQPLRPLWMBHRRASRYAKLLCGGGAAIIGKTL  
 FNMARVRLKLPILIPASQLDLSGMPVAGISGDIYHSVSRAPRPMFMCCLLLLSGV  
 GIYLLPNR"  
 330..902  
 /product="core protein"  
 903..1478  
 /product="E1"  
 1479..2756  
 /product="E2"  
 2757..3407  
 /product="NS2"  
 3408..5300  
 /product="NS3"  
 5301..5462  
 /product="NS4A"  
 5463..6245  
 /product="NS4B"  
 6246..7586  
 /product="NS5A"



mat\_peptide 7587..9359  
 /product="NS5B"  
 misc\_feature 1479..1559  
 /note="hypervariable region (HVR)"  
 misc\_feature 6954..7073  
 /note="interferon sensitivity determining region (ISDR)"

## ORIGIN

Query Match 88.3%; Score 3312.2; DB 14; Length 9410;  
 Best Local Similarity 92.7%; Pred. No. 0;  
 Matches 3476; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

1 TGGAGAGGCGCTCTTCAAGGCTCAGCCAGTGAGTCCCACTTCTGTCCCAAAAG 60  
 5001 TGGAGAGGCGCTCTTCAAGGCTCAGCCAGTGAGTCCCACTTCTGTCCCAAAAG 5060

61 CAGGACAGAGCAACTTCCCTTCTGTGAGCGTACAGGCTACTGTGTGCGTGGGCC 120  
 5061 CAGGACAGAGCAACTTCCCTTCTGTGAGCGTACAGGCTACTGTGTGCGTGGGCC 5120

121 CAGGCGCCACTCTCATGATGGAGTCAATGTGAAGTCTCATACGGCTTAAGCTTACT 180  
 5121 CAGGCTTCACCTCTCATGATGGAGTCAATGTGAAGTCTCATACGGCTTAAGCTTACT 5180

181 CTGCGGCGGCAACACCTTCTGTATAGGCTGGAGCCGTCCAAAAGAGTCACTTC 240  
 5181 CTGCAATGGGCAACACCTTCTGTATAGGCTGGAGCCGTCCAAAAGAGTCACTTC 5240

241 ACAACCCCAATCAAAATTCATATGAGTGCATGTGACCCGACTGAGAGTGTGACG 300  
 5241 ACAACCCCAATCAAAATTCATATGAGTGCATGTGACCCGACTGAGAGTGTGACG 5300

301 AGACCTGGGTGTGTGGGCGGGTCCCTTGAGCTTGTGCTGCTATTTCTTGACACA 360  
 5301 AGACCTGGGTGTGTGGGCGGGTCCCTTGAGCTTGTGCTGCTATTTCTTGACACA 5360

361 GGCAGCGTGTCTATTGTGGGTAGATCATTTGTCGGGCGGCGGCTATTGTTCCGAC 420  
 5361 GGCAGCGTGTCTATTGTGGGCAAGATCATTTGTCGGGAAAGCGGCAATTCCTCCGAC 5420

421 AGGAAATCTCTTACAGAGATTGATGAGATGGAAGTGCAGTCCCTTAC 480  
 5421 AGGAAATCTCTTACAGGAGTTGATGAGATGGAAGTGCAGTCCCTTAC 5480

481 ATGACAGAGGAAATGAGCTGCGGAGCAATTCAGCAAAAAGCGTCCGGTGTCTGAC 540  
 5481 ATGACAGAGGAAATGAGCTGCGGAGCAATTCAGCAAAAAGCGTCCGGTGTCTGAC 5540

541 ACAGCACCAAGCAAGGAGGCGGCTGCTCCGTTGGAGTCCAAATGGCGAGCCCTT 600  
 5541 ACAGCACCAAGCAAGGAGGCGGCTGCTCCGTTGGAGTCCAAATGGCGAGCCCTT 5600

601 GAGACCTTCTGGGCGAAACATATGGAATTCATACGCGGATACAGTACCTTACAGAGC 660  
 5601 GAGACCTTCTGGGCGAAACATATGGAATTCATACGCGGATACAGTACCTTACAGAGC 5660

661 TTGTCACTTGTCTGGGAAATCCCGGATGTCATCACTGATGGCTTCAAGCTCTGTCT 720  
 5661 TTGTCACTTGTCTGGGAAATCCCGGATGTCATCACTGATGGCTTCAAGCTCTGTCT 5720

721 ACTAGCCGCTCAACCAATCAACCTCTGCTTAATCAATCTGGGGGATGGATAGCC 780  
 5721 ACTAGCCGCTCAACCAATCAACCTCTGCTTAATCAATCTGGGGGATGGATAGCC 5780

781 GCCCACTGCTCCCGCAAGCTGTGCTTCACTTCTGTAGCGCCGAGCATTTGTGTGCG 840  
 5781 GCCCACTGCTCCCGCAAGCTGTGCTTCACTTCTGTAGCGCCGAGCATTTGTGTGCG 5840

841 GCTGTTGGAGCATAGGCTTGGGAAGTCTTGTGACATCTTGGGCGGCTATGAGCA 900  
 5841 GCTGTTGGAGCATAGGCTTGGGAAGTCTTGTGACATCTTGGGCGGCTATGAGCA 5900

901 GAGTGGAGGCGGCTGTGCTTAAAGTATGAGCGGCAATGCTTCCACCGAG 960

5901 GGGGTGAGAGCGCGCTCGTGCTTAAAGTATGAGCGGCAATGCTTCCCTCACAGAG 5960

961 GACCTGTTAACTTACTCCCTGCGCATCTCTCTGAGGCGCTGCTGCGGCTGCTG 1020

5961 GACCTGTTAACTTACTCCCTGCGCATCTCTCTGAGGCGCTGCTGCGGCTGCTG 6020

1021 TGGCAGAGCACTAGCTGCGGCACTGAGTCCAGAGGAGGAGGAGCTGAGTGAATGAA 1080

6021 TGGCAGAGCACTAGCTGCGGCACTGAGTCCAGAGGAGGAGGAGCTGAGTGAATGAA 6080

1081 CCGCTGATAGCGTTCGCTTCCGAGGAGTAAACATGTTTCCCAAGCACTATGTCAGAG 1140

6081 CCGCTGATAGCGTTCGCTTCCGAGGAGTAAACATGTTTCCCAAGCACTATGTCAGAG 6140

1141 AGCAGAGCGGCAAGAGTGTCACTAGATCTCTCCGAGCTTACTATACCCAACTGTG 1200

6141 AGCAGAGCGGCAAGAGTGTCACTAGATCTCTCCGAGCTTACTATACCCAACTGTG 6200

1201 AAGAGGCTCCACAGTGAATTAAGAGAGTGTCTCCAGCGCCCTGCTCCGCTGTGAGCT 1260

6201 AAGAGGCTCCACAGTGAATTAAGAGAGTGTCTCCAGCGCCCTGCTCCGCTGTGAGCT 6260

1261 AGGATGTTTGGAGTGAATATGACAGATTTTGGCTGACTTCAAGACTTGCAGTCC 1320

6261 AGGATGTTTGGAGTGAATATGACAGATTTTGGCTGACTTCAAGACTTGCAGTCC 6320

1321 AAGCTCTGCGGAGTATTCAGGAGTCCCTTTTCTCATGCAAGTGGGTAAAGAGG 1380

6321 AAGCTCTGCGGAGTATTCAGGAGTCCCTTTTCTCATGCAAGTGGGTAAAGAGG 6380

1381 GTCGAGGAGGAGAGAGGATCATGACAGACCACTGCTCATGTGAGACAGATCACCGGA 1440

6381 GTCGAGGAGGAGAGAGGATCATGACAGACCACTGCTCATGTGAGACAGATCACCGGA 6440

1441 CATGTCAAAAACGGTTCATAGAGATGTTGGCTTAAGACTGTATGATCATGTGAGAT 1500

6441 CATGTCAAAAACGGTTCATAGAGATGTTGGCTTAAGACTGTATGATCATGTGAGAT 6500

1501 GGAACATTTCCCATCAAGCATTCACAGAGGCGCCCTGCAAGCTTCCCAAGGCCAAG 1560

6501 GGAACATTTCCCATCAAGCATTCACAGAGGCGCCCTGCAAGCTTCCCAAGGCCAAG 6560

1561 TATTCCAGGCGCTGTGGGAGTGTGCTGCTGAGAGATGAGTGAAGGTTACGGGGGTGG 1620

6561 TATTCCAGGCGCTGTGGGAGTGTGCTGCTGAGAGATGAGTGAAGGTTACGGGGGTGG 6620

1621 GATTTCACCTACGTGACAGATGACCACTGACAACTGAATGACCCGTGCTCAAGTTC 1680

6621 GATTTCACCTACGTGACAGATGACCACTGACAACTGAATGACCCGTGCTCAAGTTC 6680

1681 GCCCCGAATTTCTTCAAGAGTGTGATGAGGCTGCAAGTGAAGTGAAGTGAAGTGA 1740

6681 GCCCCGAATTTCTTCAAGAGTGTGATGAGGCTGCAAGTGAAGTGAAGTGAAGTGA 6740

1741 AAACTCTCTTACAGGAGATGATTCAGAGTGGCTGCTCAAGTGAAGTGAAGTGAAG 1800

6741 AAACTCTCTTACAGGAGATGATTCAGAGTGGCTGCTCAAGTGAAGTGAAGTGAAG 6800

1801 TCGCAGCTCCATGACAGCCGAAACGAGTATGAGAGTCTCACTTCAATGCTCACGAG 1860

6801 TCGCAGCTCCATGACAGCCGAAACGAGTATGAGAGTCTCACTTCAATGCTCACGAG 6860

1861 CCTTCCACATCAACAGAGAGCGGTTAAGCGAGGCTGACAGGAGGCTTCCCTCC 1920

6861 CCTTCCACATCAACAGAGAGCGGTTAAGCGAGGCTGACAGGAGGCTTCCCTCC 6920

1921 TTGGCAGGCTTCACTAGCTGAGTGTGCTGAGGCTTCTTCAAGGCGCATATGATTAAC 1980

6921 TTGGCAGGCTTCACTAGCTGAGTGTGCTGAGGCTTCTTCAAGGCGCATATGATTAAC 6980

1981 CAAATGACTTCCACAGCTGACTCATGAGGCCAATCTCTGTGGCGGATGAGATG 2040

Db 6981 CGTCAAGACTCCCAAGATGCGCACTCATATGAGGCCAACCTCTGTGGCGGAGAGATG 7040  
 QY 2041 GGGGGGAGATTACCCCGGTGAGTCAAGAAACAAGTAGTAATCTTGACCTTTTCAC 2100  
 Db 7041 GGGGGGAGATTACCCCGGTGAGTCAAGAAACAAGTAGTAATCTTGACCTTTTCAC 7100  
 QY 2101 CCGCTCCGAGCGGAGAGATGAGCGGAGATGTCCGTCCGCGGAGATCTTGCGGAA 2160  
 Db 7101 CCGCTCCGAGCGGAGAGATGAGCGGAGATGTCCGTCCGCGGAGATCTTGCGGAA 7160  
 QY 2161 TCCAAAGAAATTCACCAAGAGATGCGGAGATGCGGAGATTAACAACCTTCGCG 2220  
 Db 7161 TCCAAAGAAATTCACCAAGAGATGCGGAGATGCGGAGATTAACAACCTTCGCG 7220  
 QY 2221 CTGAGGCTCGGAAAGCGCGGAGTACGTCTCCAGTGATCATGGGTGCGGAGTCC 2280  
 Db 7221 ATGAGATCTGGAAGAGACCGGAGTACGTCTCCAGTGATCATGGGTGCGGAGTCC 7280  
 QY 2281 CCTACTAAGACCCCTCTATATACCACTCAAGAGAGAGAGAGAGTGTCTGAGAGA 2340  
 Db 7281 CCTACCAAGGCCCTCTCAATATACCACTCAAGAGAGAGAGAGTGTCTGAGAGA 7340  
 QY 2341 TCCACCGTGTCTTCTGCGGTGAGAGTCTGCAACAAGGCTTTCGTAGCTCCGAA 2400  
 Db 7341 TCCACCGTGTCTTCTGCGGTGAGAGTCTGCAACAAGGCTTTCGTAGCTCCGAA 7400  
 QY 2401 TCGGCGGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
 Db 7401 TCGGCGGTGAG 7460  
 QY 2461 GCAAGATCTGAGAGTGTATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
 Db 7461 GCAAGATCTGAGAGTGTATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7520  
 QY 2521 CCGGATCTCAAG 2580  
 Db 7521 CCGGATCTCAAG 7580  
 QY 2581 TGTGCTGATGATCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
 Db 7581 TGTGCTGATGATCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7640  
 QY 2641 AGCAAGTGTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
 Db 7641 AGCAAGTGTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7700  
 QY 2701 GCTACCAATCCGAG 2760  
 Db 7701 GCTACCAATCCGAG 7760  
 QY 2761 ATCTGAGAGATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
 Db 7761 ATCTGAGAGATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7820  
 QY 2821 AAGGCTAAGTCTTATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
 Db 7821 AAGGCTAAGTCTTATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7880  
 QY 2881 TCTAATTTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
 Db 7881 TCTAATTTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7940  
 QY 2941 ATCCGCTCCGTGTGAG 3000  
 Db 7941 ATCCGCTCCGTGTGAG 8000  
 QY 3001 ATGGCAAAATGAGTGTCTGAGTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
 Db 8001 ATGGCAAAATGAGTGTCTGAGTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8060  
 QY 3061 CTATGCTGTGAG 3120  
 Db 8061 CTATGCTGTGAG 8120

QY 3121 GTCTCACCCCTCTCAAG 3180  
 Db 8121 GTCTCACCCCTCTCAAG 8180  
 QY 3181 CAGCGGTGTGAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
 Db 8181 CAGCGGTGTGAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8240  
 QY 3241 TATGACACCCGCTGTGTGAGTCAACAGTCACTGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
 Db 8241 TATGACACCCGCTGTGTGAGTCAACAGTCACTGAGAGAGAGAGAGAGAGAGAGAGAG 8300  
 QY 3301 ATTATCAATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3360  
 Db 8301 ATTATCAATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 8360  
 QY 3361 CCGCTTATATGAG 3420  
 Db 8361 CCGCTTATATGAG 8420  
 QY 3421 TGCAGCGAG 3480  
 Db 8421 TGCAGCGAG 8480  
 QY 3481 GCCTGTGAG 3540  
 Db 8481 GCCTGTGAG 8540  
 QY 3541 GGCCTGTGTGATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
 Db 8541 GGCCTGTGTGATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8600  
 QY 3601 TTACAG 3660  
 Db 8601 TTACAG 8660  
 QY 3661 GACCTGTGAGTGTGATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720  
 Db 8661 GACCTGTGAGTGTGATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8720  
 QY 3721 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
 Db 8721 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8820  
 QY 3781 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8880  
 Db 8881 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8940  
 QY 3941 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9000  
 Db 9001 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9060  
 QY 9061 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9120  
 Db 9121 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9180  
 QY 9181 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9240  
 Db 9241 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9300  
 QY 9301 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9360  
 Db 9361 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9420  
 QY 9421 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9480  
 Db 9481 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9540  
 QY 9541 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9600  
 Db 9601 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9660  
 QY 9661 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9720  
 Db 9721 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9780  
 QY 9781 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9840  
 Db 9841 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9900  
 QY 9901 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9960  
 Db 9961 AAAAGGTATGATCACTCAACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10000

## RESULT 6

## HPCKLR2

## LOCUS

## DEFINITION

HPCKLR2 9410 bp RNA linear VRL 10-FRB-1999  
 Hepatitis C virus (strain HCV-1b, clone HCV-K1-R2), complete genome  
 sequence.

## ACCESSION

## VERSION

## KEYWORDS

D50481.1 GI:1030705  
 polypeptide; core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A;  
 NS5B; envelope protein; non-structural protein;  
 interferon-sensitive; interferon-resistant; IFN-sensitive;  
 IFN-resistant; ISDR; interferon sensitivity determining region;  
 HVR; hypervariable region.

## SOURCE

## ORGANISM

Hepatitis C virus  
 Hepatitis C virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.  
 1 (sites)

## REFERENCE

## AUTHORS

Yamamoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T.,  
 Yamamoto, C., Izumi, N., Marumo, F. and Sato, C.  
 Comparison of full-length sequences of interferon-sensitive and  
 resistant hepatitis C virus 1b. Sensitivity to interferon is  
 conferred by amino acid substitutions in the NS5A region  
 J. Clin. Invest. 96 (1), 224-230 (1995)  
 MEDLINE  
 PUBMED  
 7542279  
 2 (bases 1 to 9410)  
 REFERENCES  
 Enomoto, N.

JOURNAL REFERENCE AUTHORS TITLE JOURNAL	Unpublished 3 (bases 1 to 9410) Enomoto, N. Direct Submission Submitted (08-MAY-1994)
---	---

QY	661	TTGTGCACTGCGCCGGGAATCCGCGATTGCACTGAAGGGGTTCAACGCTCTGTC	720
Db	5661	TTGTTCACACTGCGCTGGGAACCCGCGATPACATATTGAAGGATTCACAGCTTCATC	5720
QY	721	ACTAGCCGCGTCACCAACCCAATTACCTCTCGTCTTAACTCTGGGGGATGGGTAGCC	780
Db	5721	AACGACCCGCTCACCAACCCAACACTCTCTGTTTAAATATCTTGGAGATGGGTGGCC	5780
QY	781	GCCCAACTGCTCCCCCAAGTGTGCTTCAAGCTTTGTAAGCGCCGCGCATTTGTGTGCG	840
Db	5781	GCCCAACTGCTCCCCCAAGAGTGCCTCGGCGCTTCGTAAGCGCGTGCATGCGCGGTGCG	5840
QY	841	GCTGTGGCAGATGAGGCGCTTGGGAAAGGTGCTGTGGACATCTTGGCGGGCTATGGAGCA	900
Db	5841	GCTGTGTGACAGATAGGCGCTTGGGAAAGGTGCTGTGGACATCTTGGCGGGCTTATGGAGCA	5900
QY	901	GGAATGGACAGGCGCGCTGTGGGCTTTTAAGGTCAATGACGGCGGAATGCCCTCCACAGAG	960
Db	5901	GGGGTGGACAGGCGCGCTGTGGGCTTTCAAAGTCAATGACGGCGGATATGCCCTCCACGAG	5960
QY	961	GACCTGTGTAATTACTCCCTGTCATCTCTCTCTCTGCTGTCGCGGTGTCGCGGGGTGCTG	1020
Db	5961	GACCTGTGTAATTACTCCCTGTCATCTCTCTCTCTGCTGTCGCGGTGTCGCGGGGTGCTG	6020
QY	1021	TGCGGAGGGAATATCTGCGCTGCGGACGTGGGCTCAGAGGGAGGGGGCTGTGCAATGAC	1080
Db	6021	TGCGGAGGGAATATCTGCGCTGCGGATGTGGGCGCGGGAGGGGGCTGTGCAATGATAC	6080
QY	1081	CGGCTGATAGCGTTGGCCCTCGCGGGGGTAAACATGTTTTCCCGACGCACTATGTGSCAGAG	1140
Db	6081	CGGCTGATAGCGTTGGCTTGGCTTGGGGGGCAACAGCTCTCCCGACGCACTATGTGCTGAG	6140
QY	1141	AGCGGCGCGGACGACAGTGTCACTCAGATCTCTCGACCTTACTATCACCCACTGTG	1200
Db	6141	AGCGGCGCGGACGCGCGGTGTCATCTCAGATCTCTCAGCTTATCATCATCACTGATG	6200
QY	1201	AAGAGGCTTCACCAATGATTAACAGAGATGCTCTCAAGCCCTGCTCGGCTCGTGCTA	1260
Db	6201	AAGAGGCTTCACCAATGATTAATAGGATTTGCTTCAAGCCATGCTCGGCTCGTGCTA	6260
QY	1261	AGGAGTGTGGGACGTGATATGCAAGTTTTGGTGACTTCAAGACCTGGCTCAAGTCC	1320
Db	6261	AGGAGTGTGGGAGCTGATATGCAAGGTGCTGTGTGACTTCAAGACTTGGCTCMAATCC	6320
QY	1321	AAAGCTCTGCGCGGATTAACGGGAGTCCCTTTTTCATGCAACGTGGGTACAAGGGG	1380
Db	6321	AAAGTCTGCGCGGCTGCGAGAGATCCCTTTTCTCATGCAACGTGGGTACAAGGGA	6380
QY	1381	GTTCTGGCGGGAGACGGCATATGCAAGCAACCTGCTCATGTGGAGCAAGATCACGGGA	1440
Db	6381	GTTCTGGCGGGAGACGGCATATGCAAGCAACTTGCCTCATGTGGAGCAAGATCACGGGA	6440
QY	1441	CATGTCAAAAAGGTTCCATGAGATGCTGTGGGCTTAAGCCTGTAGTAATGTGGCAT	1500
Db	6441	CATGTCAAGAGGTTCCATGAGAAATCGTTGGGCTTAAGACTGTGAGCAATACGTGGCAT	6500
QY	1501	GGAAGATTTCCCATCAACGATATACCAACGAGGCCCTGACAGCGCTCCCGACGCGCAAC	1560
Db	6501	GGAAGATTTCCCATCAACGCGATACCAACGAGGCCCTGCAACACTCCCGGCGCGCAAC	6560
QY	1561	TATTCAGGGCGCTGTGGCGGGTGTGCTGTGAAGTACGTAGAGTTTACGCGGGTGGGG	1620
Db	6561	TATTCAGAGGCGCTGTGGCGGGTGTGCTGTGAAGAGTACGTAGAGTTTACCGGGTGGGG	6620
QY	1621	GATTTTCCACTAGTACGAGCATGACACTGACCAACGTAAAAATGGCCGTGCAAGTTTCCA	1680
Db	6621	GATTTTCCACTAGTACGAGCATGACACTGACCAACTAAAAATGCCGTGCAAGTTTCCG	6680
QY	1681	GCCCCCGAATTTCTTCAAGATGATGGGAGGCGGCTGACAGGTAGCGCTTCGGCGTGC	1740
Db	6681	GCCCCCGAATTTCTTCAAGAAATCTGAGCGGGAGTGGCGGTGTCACAGGTATCGCTTCGGCGTGC	6740
QY	1741	AAACCTCTTCAACGGAGAGGTCACTTCCAGGTCGAGGCTTCAACCAATACCTGTGTTGG	1800

Db	6741	AAACTCTCTCGCGGGATGAGGTCACTTCCAGGTCGGGCTCAACAAATAGTGTGGA	6800
Qy	1801	TCGAGCTCCCATGCGAGCCCGAAACGAGATGACAGTCTCACTTCATGCTCACCGAC	1860
Db	6801	TCACAGCTCCCATGCGAGCCCGAAACGAGATGACAGTCTCACTTCATGCTCACCGAC	6860
Qy	1861	CCCTCCCAATCAACAGCAGAGACGGCTAAGCGGAGGCTGGCCAGGGGGTCTCCCCCTCC	1920
Db	6861	CCCTCCCAATCAACAGCAGAGACGGCTAAGCGGAGGCTGGCCAGGGGGTCTCCCCCTCC	6920
Qy	1921	TTGGCCAGCTTTTCACTAGTAGCCAACTTGTCTGCGCTTCTCGAAGGGACATACATTACC	1980
Db	6921	TTGGCCAGCTTTTCACTAGTAGCCAACTTGTCTGCGCTTCTCGAAGGGACATACATTACC	6980
Qy	1981	CAAAATGACTTCCGAGAGCTGACCTTCATGAGGCCAACCTCTGTGGCGCATGAGATG	2040
Db	6981	CATATGACTTCCGAGAGCCGACCTTCATGAGGCCAACCTCTGTGGCGCATGAGATG	7040
Qy	2041	GGCGGGGACATTAACCCGCTGAGATCAAGAAACAAGTAGTAATCTTGAGCTTTTGAC	2100
Db	7041	GGCGGGGACATTAACCCGCTGAGATCAAGAAACAAGTAGTAATCTTGAGCTTTTGAG	7100
Qy	2101	CGGCTCCGAGGGAGAGAGATGACCGGGAATGTCCGCTCCGCGGGAGATCTCTCGGAAA	2160
Db	7101	CGGCTTCCAGCGGAGAGAGATGAGAGGGAGTATCCCTTCCGCGGAGATCTCTCGAAA	7160
Qy	2161	TCCAAAATTCGACACGAGCATCCCGCATGGGACGCGCGGATTAACAACCTTCGCTG	2220
Db	7161	TCGAGAAATTCGCGCGGCAATGCCCATATGGGACGCGCGGATTAACAACCTTCACCTG	7220
Qy	2221	CTGAGTCTTGAAAGGCCCGGACATAGTCTCTTCCATGAGTACATGGGTGCCCATGCCA	2280
Db	7221	ATAGAGTCTTGAAAGGCCCGGACATAGTCTCTTCCATGAGTACATGGGTGCCCATGCCA	7280
Qy	2281	CCTATTAAGACCCCTCTTAACAACCTCAAGAGAGAGACAGTTGTTCTGACAGAA	2340
Db	7281	CCTACAAAGGCCCTCTCAATACACCTTCAAGAGAGAGAGAGAGTGTCTTCAAGAA	7340
Qy	2341	TCACCGTGTCTTCTGCCCCGCGAGCTTGCCCAAAAGCTTTCCGTAAGCTTCGAAACCG	2400
Db	7341	TCACGTGTCTTCTGCCCCGCGAGCTTGCTCAAAAGACCTTCGAGAGCTTCGAAATCG	7400
Qy	2401	TCGCGCTGCAACGCGGACACGCGACCGCCCTCTCTGAACAACCTTCGACATGGCGGA	2460
Db	7401	TCGCGCGCGATAGCGGACACGCGACCGCCCTCTCTGAACAACCTTCGAGGACGGCGAC	7460
Qy	2461	GCAGGATCTGACGTTGAGTGCCTTAATCCCTCAATGCCGCCCTTGAGGGGGAGCCGGGGAC	2520
Db	7461	GCAGGATCCGAGTTGAGTGCCTTAATCCCTCAATGCCGCCCTTGAGGGGGAGCCGGGGAC	7520
Qy	2521	CCCGATCTCACGACGAGGCTTTTGGTCTTACCGTGAATGAGAGGCGCGGTGAGAGCGTGC	2580
Db	7521	CCCGACCTCAAGAGAGGCTTTTGGTCTTACCGTGAATGAGAGGCGCGGTGAGAGCGTGC	7580
Qy	2581	TGCTGCTGATGTCTTACATATGACACAGGCGCTCTGATCACGCCATGGCTCTCGAGGAA	2640
Db	7581	TGCTGCTGATGTCTTACATATGACACAGGCGCTCTGATCACGCCATGGCTCTCGAGGAA	7640
Qy	2641	AGCAAGCTGCCCATCAACGCGTTTGAAGAACTTTTGTGCGGTACCAACATAGGTGAC	2700
Db	7641	AGCAAGTTACCATCAATGAAGCGTTTGAAGAACTTTTGTGCGGTACCAATATAGGTGAT	7700
Qy	2701	GCTACACATCCCGAGCGCAAGCCAGCGGACGAGAAAGTCACTTTGACAGCTGCAA	2760
Db	7701	GCTACACATCCCGAGCGCAAGCCAGCGGACGAGAAAGTCACTTTGACAGCTGCAA	7760
Qy	2761	ATCTTGAAGATCACTTACACAGACGCTGCTCAAGAGATGAAGGCGGAAGCGTCAAGTT	2820
Db	7761	GTCTTGAAGACACTTACCGGGAAGTGTCTCAAGGAGATGAAGGCGGAAGCGTCAAGTT	7820
Qy	2821	AAAGCTAAGCTTATCACTAGAGAGGCTTGCAGAGCTGACGCCCCCAATTGGCCAAA	2880

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 9373)	Negayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and Sato, C.	Characteristics of hepatitis C viral genome associated with disease progression	Unpublished				
2 (bases 1 to 9373)	Negayama, K., Kurosaki, M., Enomoto, N., Izumi, N. and Sato, C.	Direct Submision	Submitted (24-NOV-1999)				
Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan							
FEATURES							
SOURCE							
1. .9373	Location/Qualifiers						
/organism="Hepatitis C virus"							
/mol_type="genomic RNA"							
/strain="MDJ4"							
/db_xref="taxon:11103"							
/note="from a patient with persistently normal ALT values (ASC-14)"							
1. .9373							
/gene="MDJ4"							
330. .9356							
/gene="MDJ4"							
/codon_start=1							
/product="polyprotein"							
/protein_id="A061205.1"							
/db_xref="GI:7341103"							
/translation="MSTNPKDKRKTNRNRPRDQYKPGGSGOIVGGVILIPRGPRLGVATRTKTSERSQPRGRQRIPKARQDEGAAMQPGVPMWLYENEGMAGWLLSPGSRNMGFTDPRRRSRNDKGVLDITGCFALWKYITLVGAPLVNAGAAALAHGVRLEDSGVNATGNLPGCSFSIFLLALLSLTPASAYVRNVSQVYHTINDCSNSIYEADMINHSGCVCEVREGVSRVCVATPTLARNCSVPTTIRRVLDLGAACASAVYVGLCSQFLVLSQLTFFSPRRHETVODNCISIVGLSHGRAMDMLNSPTAALVYVGLSRIPOAMVMDVAGAHGCVTLAGLAVYSGVMAKYLIMLLPAGDGMWLLIGSSQDTRTRFSLSPSPGSONIQLVKTGSMHINRTALNCDSHOTGTLLALFAHKINTSSCLESLLTCTIDKFDQCGMGLTVYEHGSDSKPCWMAHPKCIIVASKVCQVPCYCTPSVVVGTGTDPRFAPPTWGENETDVLNNTNRPDGMNCGVMNSGGYTCGCPPTCNICGVNDTLICTPDCFRKHPRPATYKCGSGWMLIPRCWVDYPRLMHPCVNFVFKRYMAGVGEHRLSAACWMTGRBCNLBDRSELSPLLSTTEMQIIPCSFTLPAI LSGILHINQIVDQYLVGVGASVPLVKKMEVYLILLIADADICLIMMLIIPQAPRLBNLVVLAASVAGAHGLSLFLVFCAMAYIKGALPGCAALYVWVPLILLLLALPAPRAYMDREMAASCAGAVFGLVTLSPYVYVFLRLIMWSQITVTRAEHLQVWVPLNVNRGGDAIILITCAVHRELFDITKLILILGVLNLOASIKVYFVRVQGLIRACMLARKVAGGVQMAFMKLAALNTQYVDHLTPQDMAHSLRLAVAVEVQVSDMEKLIITWGLDTRAAAGDITLGLPVSAARKELLIPGADLEGGKRLALPITYASOOTRGLGCTITSLTGRDNQVGEVQVSTQSFALCNVNGCTVPHGASGKTLAAGKPGITQMTNVDDQLVGMQAPPGARSLTPCTGSSDLVTRHADVLVPRRGDSRGLSLPRPVLYKSGSGGGLPCPSHAGVIFPAACVTRGVAAVDFVPVSMETLFRSPFTNNSPPVAPQTFQVNAHLHAPGSGKSTVPAAYAAQYKVLVINSVATLQFSAVNSKAHGVDPNIRTGARVTVTGAPITVSTYKGLFADGGCGGAVIDIMCDSHSDSTIILIGTGLDQAEFRAGARLVVLAATPAGSVTVHPNIEVALNTEGIIPYGKAIIPITIKGGRHLITCSKKKCDLALYLSGLINAAVYAGLDVSVYPASGVNVVADALMTQVTFGDSVIDNCTQTVDFSLDFTFTEITTVQDAVSSRSQKRGRRGARGIATVYTPGERPSGMDSSVLYLCECYDAGCAWELPTAETSIVRAALNTGLFVPCDHLFEWESVFTGLTHILDAHFLSQTQAGDNFPLYALVQYCARAAQPPSDQWKKICILRLPETHLGPPTLIRLGAVONEVTLITLPIKYIMACASDALEVTSTVNVGVLALLAAYCLITGTSVIVIGRIILISGKPAVIDBRDLVYAEPEMCEASLPIYBEGMLAIOFKQKALGILQTRATQAEVAAVYESKRALBAPAKAMMISGITYLALGSLTPGNDPITABLMFVATISPLTTSSTLNFNLILGGWAAALAPKASAVAGAGIAGAASITGLKGLVILLAGVAGAGALVAFKMSGENPSTDLNLRPAIILSNALVGVCAAILLRHVGEAGVQMMNRLLIAFASRGHNSPHYVPSBDAAARVQIISLITLQLLKLVHSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGMHTDQVPCQVPAEPFETVDUGVRLHYAPACKPLPESVTOQELNOLVGSQVLPCEBEPVAVLTLDPSTITABAKRLARGSPSPSLASSASQLSAPSLKATCTTHSDSDALIINLMLRQMGENTITRVSEKRLVILSDGDLPLTAEDEREVPAIILAKSKRPAMPIMVARDPNPLLSBQMDVYVVIDGPELPTAEKAPIPPPRKRTVLTSTVSALAEATKPTDSSGSAVDGATAPPDQSDNAGSDIVESVSMPLGEGQDPLITGPKTCSMTMGTPPINAVTGPGSPAPNYERBALMRVAAEYVETRVGDPHYVTGM							

VKDLIEDTETPIDTTIMAKNEIFCVQPERGKRPRLIVFDPDLGVRCEKNALYDVV  
STLPQAVMSSYXFOYSPGQRYVEFLVDAMSKCPMGAFRCPSFTYREDIRIEE  
S1YOCDDLAPEAROALIRSLIBERLYVGGLTNSKNGCGRGACSVLTGSGNTITC  
YUKASAAACRAKLRDCTMLVCGDDLVIBESACTOEDPASLAVFEMAMRYAPRDP  
POPEYDLIELITSSSNVSAVHADSGKRVYTLTRDPTPLARAMEFARSTPVNSIGN  
IIMYADTLMARMLIMTHFPFSLIAQDLKALDQIYGACYSLEPDLQIILQRHLG  
SAFSLHSYSPGEINRYASCIRKGVPLVWHRARSAVAKLSIQGRATAGSKYLFN  
MAVRTKLKLPIPIPAASQDLISNMFVAGSGDIYHSLSBARPRWFMLCLLLSVGVI  
YLLPNR"

ORIGIN

Query Match 87.9%; Score 3297.8; DB 14; Length 9373;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 3467; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

QY 1 TGGGAGGGCGCTTTCACAGGCTTCACCCAGTGGATGCCCACTTCCTGCCCAAGG 60  
DB TGGGAGGGCGCTTTCACAGGCTTCACCCAGTGGATGCCCACTTCCTGCCCAAGG 5054  
QY 61 CAGGACGAGAGACAATTCCCTACTGTGGCGTACAGGCTACTGTGTGCGCTAGGGCC 120  
DB CAGGACGAGAGATACTTCCCTACTGTGGCGTACAGGCTACTGTGTGCGCGCAAGGCT 5114  
QY 5055 CAGGACGAGAGATACTTCCCTACTGTGGCGTACAGGCTACTGTGTGCGCGCAAGGCT 5114  
DB 121 CAGGCGCCACCTTCATCATGAGATCAATGTGGAAGTGTCTCATACGGCTTAAAGCTTACT 180  
QY 5115 CAGGCTCAACCCCATGTGGATCAAAATGTGGAAGTGTCTCATACGGCTTAAAGCTTACG 5174  
DB 181 CTGGGGGGGGCAACCCCTGTGTATAGGCTGGGAGCGCTCCAAAACGAGGTCACCGTC 240  
QY 5175 CTACAGGGGCCAACGCCCCGTGTATAGGCTTAGGAGCCGTCCAAAATGAGTCACCGTC 5234  
DB 241 ACACACCCCATTAACAAATTCATCATGAGATGAGATGAGCCGACCTGAGAGTCTGACG 300  
QY 5235 ACACACCCCATTAACAAATTCATCATGAGATGAGATGAGCCGACCTGAGAGTCTGACG 5294  
DB 301 AGCACTGGGTGTGTGTGGGGGGGTCCTTTCAGAGCTGTGGCTGTGATTCCTTGACAACA 360  
QY 5295 AGCACTGGGTGTGTGTGGGGGGGTCCTTTCAGAGCTGTGGCTGTGATTCCTTGACAACA 5354  
DB 361 GGCAGCGTGTCTATGTGGGTAGATTCATTCCTGCGGGGGCGCGGTATTTGTTCCCGAC 420  
QY 5355 GGCAGCGTGTCTATGTGGGTAGATTCATTCCTGCGGGGGCGCGGTATTTGTTCCCGAC 5414  
DB 421 AGGAGATCTCTTACAGAGATTCGATGAGATGGAAGATGCGGCTGCACTCCCTTAC 480  
QY 5415 AGGAGATCTCTTATCGGGAGTTGATGATGAGAAAGTCCGATGCGACCTCCCTTAC 5474  
DB 481 ATCGAGCAGGGAATGAGCTTCGCCGAGCAGTTCAGAGAAAAGCGCTCGGGTGTGTCGAG 540  
QY 5475 ATCGAGCAGGGAATGAGCTTCGCCGAGCAGTTCAGAGAAAAGCGCTCGGGTGTGTCGAA 5534  
DB 541 ACAGCCACCAAGCAGGAGGCGCGTGTCTCCGTGTGAGATTCAGATGCGGACCGCTT 600  
QY 5535 ACAGCCACCAAGCAGGAGGAGTTCGCTCCGTGTGAGATTCAGATGCGGACCGCTT 5594  
DB 601 GAGACTTCTTGGGGGGAACACATGTGGAATCTTATCAGCGGGAATACAGTACTTAGAGAGC 660  
QY 5595 GAGGCTTTTGGGGGGAACACATGTGGAATCTTATCAGCGGGAATACAGTACTTAGAGAGC 5654  
DB 661 TTGTCACTGTGCTGGGAATTCGCCGATGTCATGATGAGGCTTCAAGCTCTGTGC 720  
QY 5655 TTGTCACTGTGCTGGGAATTCGCCGATGTCATGATGAGGCTTCAAGCTCTGTGC 5714  
DB 721 ACTAGCCCGCTCACACCACTTACCTCTGTCTTAAATCTCTGGGGGAGATGGGTAGCC 780  
QY 5715 ACCAGCCCGCTCACACCACTTACCTCTGTCTTAAATCTCTGGGGGAGATGGGTAGCC 5774  
DB 781 GCCCAACTGCTCCCGCAAGTGTGCTTACGCTTTCGTAAGGCGCGGCAATTCGCTGAGGC 840  
QY 5775 GCCCAACTGCTCCCGCAAGTGTGCTTTCGTAAGGCGCGGCAATTCGCTGAGGC 5834  
DB 841 GCTGTGGCAGCATAGGCTTTGGGAAGGTGCTTGTGACATCTTGGCGGGCTATGAGCA 900

DB 5835 GCTGTGGCAGCATAGGCTTTGGGAAGGTGCTTGTGACATCTTGGCGGGCTATGAGGCA 5894  
QY 901 GGAATGGCAGGCGCGCTCGTGGCTTTTAAAGTCATAGCGCGGAATGCCCTCCACCGAG 960  
DB 5895 GGGGTGGCGGGGCACTCGTGGCTTTTAAAGTCATAGAGTGGGAGATGCCCTCCACTAG 5954  
QY 961 GACCTGTGTTAACTTACTCCCTGACATCTCTCTCTGTGTCCTGTGGTGGTGGGTCTGTG 1020  
DB 5955 GACCTGTGTTAACTTACTCCCTGACATCTCTCTCTGTGTCCTGTGGTGGTGGGTGGT 6014  
QY 1021 TGGCAGCGCATCTGCGTGGAGCGTGGGTCCAGGGGAGGGGGGCTGTGAGTGGATGAAC 1080  
DB 6015 TGGCAGCGCATCTGCGGCGGATGCGGATGCGGCGGAGAGGGGGCTGTGAGTGGATGAAC 6074  
QY 1081 CGGCTGATAGCTTGGCTCGCGGCTCGCGGGATACATGTTTCCCGCAGCACTATGTGCAGAG 1140  
DB 6075 CGGCTGATAGCATTTGGCTTGGGGGATACAGTCTCCCGCAGCACTATGTGCTGAG 6134  
QY 1141 AGCGAGCGCGCAGCAGTGTCTCATAGATCTCTCTCGACCTTATCTACCCACTGTTCG 1200  
DB 6135 AGCGAGCGTGGCGGCGGCTGACCCAAATCTCTCGAGCTTATCATCATCTCATGTTTGG 6194  
QY 1201 AAGAGCTTCACACAGTGAATTAACAGAGATCTCTCCAGCTTACCTTACCTTACCTTTCG 1260  
DB 6195 AAGAGCTTCATCATGATGATCATAGAGATCTCTCCAGCTTACCTTACCTTACCTTTCG 6254  
QY 1261 AAGGATGTTTGGGATCTGATATGACAAGTTTTGGCTGACTTCAAGACTTGGCTTCCAGTCC 1320  
DB 6255 AAGGATGTTTGGGATCTGATATGACAAGTGTGTTGACTGATTTCAAGACTTGGCTTCCAGTCC 6314  
QY 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGTCACAGCTGTGGATCAAGGGG 1380  
DB 6315 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGTCACAGCTGTGGATCAAGGGG 6374  
QY 1381 GTCTGGCGGGGAGACGGGATCATGAGACACACTGCTCATGTGGAGCAGATPACCGGA 1440  
DB 6375 GTCTGGCGGGGAGACGGGATCATGAGACACACTGCTCATGTGGAGCAGATPACCGGA 6434  
QY 1441 CATGTCAAAAACGTTCCATGAGGATCGTTGGGCTTAAGACCTGTATGATTAATGTGGCAT 1500  
DB 6435 CATGTCAAAAACGTTCCATGAGGATCGTTGGGCTTAAGACCTGTATGATTAATGTGGCAT 6494  
QY 1501 GGAACATTTCCCATCAAGCATPACACACGCGGCCCTTGCACGCGCTCCCGAGGCCAAC 1560  
DB 6495 GGAACATTTCCCATCAAGCATPACACACGCGGCCCTTGCACACCTCCCGAGGCCAAC 6554  
QY 1561 TATTCAGAGGCGCTGTGGGGGTGCTGCTGAGAGATGATGAGGTTTACGGGGGTGGG 1620  
DB 6555 TATTCAGAGGCGCTGTGGGGGTGCTGCTGAGAGATGATGAGGTTTACGGGGGTGGG 6614  
QY 1621 GATTTCCATCACTGTCAGAGATGACATCTGACAACTGTAATAATGCCCGTGCAGTTTCCA 1680  
DB 6615 GATTTCCATCACTGTCAGAGATGACATCTGACAACTGTAATAATGCCCGTGCAGTTTCCG 6674  
QY 1681 GCCCCGGAATTTCTTCAAGAAATGATGAGGAGTGGGCTGTCACAGTACGCTCCGCGTGC 1740  
DB 6675 GCCCCGGAATTTCTTCAAGAAATGATGAGGAGTGGGCTGTCACAGTACGCTCCGCGTGC 6734  
QY 1741 AAACCTCTCTTACGGGAGAGTGCATATTCAGAGTGGGCTGCAACCAATACCTGTTGGG 1800  
DB 6735 AAACCTCTCTTACGGGAGAGTGCATATTCAGAGTGGGCTGCAACCAATACCTGTTGGG 6794  
QY 1801 TGCAGCTCCCATGAGGCGCGAACCGGATGATGAGTGTGCTCATCTTCCATGCTCACGAC 1860  
DB 6795 TGCAGCTCCCATGAGGCGCGAACCGGATGATGAGTGTGCTCATCTTCCATGCTCACGAC 6854  
QY 1861 CCCTCCCAATYACACAGAGACGAGTAAAGCGAGGCTGAGCGAGGGGTCTCCCTCCCTCC 1920  
DB 6855 CCCTCCCAATYACACAGAGGCGGCTAAAGCGTAAAGCGTAAAGCGGAGGTCTCCCTCCCTCC 6914  
QY 1921 TTGGCAGGCTCTTACAGTACGCAATGTTGTCTGCGCTTCTCTGAAAGCGACATTCATTAAC 1980  
DB 6915 TTGGCAGGCTCTTACAGTACGCAATGTTGTCTGCGCTTCTTGAAGCGACATTCATTAAC 6974



QY 1981 CAAATGACTTCCAGACGCTGACCTCATCGAGGCCAACCTCCTGTGGCGGCAATGAAATG 2040  
 DB 6975 CATCATGACTCTCCCGAGCGCTGACCTCATGAGGCCAACCTCCTGTGGCGGCAAGATG 7034  
 QY 2041 GGGGGGAGCAATTAACCGCGTGAAGTCAAGAAACAGGTAGTAATCTTGAATCTTTTGAC 2100  
 DB 7035 GGGGAAACATCAACCGCGTGAAGTCAAGAAATAGGTAGTAATCTTGAATCTTTTGAC 7094  
 QY 2101 CCGCTCCGAGCGGAGAGATGAGCGGAAAGTGTCCCTCCCGGAGATCTTCCGAAA 2160  
 DB 7095 CCGCTCCGAGCGGAGAGATGAGCGGAAAGTGTCCCTCCCGGAGATCTTCCGAAA 7154  
 QY 2161 TCCAAAGAAATTCACAGAGATGCCCGGAGCGGAGATTAACAACCTCCGCTG 2220  
 DB 7155 TCCAGAAATTTCCCGGAGATGCCCGGAGCGGAGATTAACAACCTCCGCTG 7214  
 QY 2221 CTGAGTCTGAGAGCGCGGAGCTAGTCCCTCAGTGTAGCAATGAGTCCCACTGCCA 2280  
 DB 7215 CTGAGTCTGAGAGCGCGGAGCTAGTCCCTCAGTGTAGCAATGAGTCCCACTGCCA 7274  
 QY 2281 CTTACTTAAGACCTCTCTATACCACTTCCAGAGAGAGACAGTGTCTTGAACAGAA 2340  
 DB 7275 CTTACCAAGGCGCTCCCAATACCACTTCCAGAGAGAGACAGTGTCTTGAACAGAA 7334  
 QY 2341 TCCACGCTGCTTCTGCGCTGGGAGCTTGCCAAAGGCTTCCGCTAGCTCCGAAACG 2400  
 DB 7335 TCCACGCTGCTTCTGCGCTGGGAGCTTGCCAAAGGCTTCCGCTAGCTCCGAAACG 7394  
 QY 2401 TCGGCGCTGAGAGCGGAGCGGAAACCGCTCTGAGCAACCTTCCAGAGAGAGAG 2460  
 DB 7395 TCGGCGCTGAGAGCGGAGCGGAAACCGCTCTGAGCAACCTTCCAGAGAGAGAG 7454  
 QY 2461 GCAAGATCTGAGCTGAGTGTATCTCTCATGCCCCCTTGAAGGAGAGCGGAGAG 2520  
 DB 7455 GCAAGATCTGAGCTGAGTGTATCTCTCATGCCCCCTTGAAGGAGAGCGGAGAG 7514  
 QY 2521 CCGGATCTGAGAGCGGAGCTTGTGTACCTGAGTGAAGAGCGGAGAGAGAGTGTGTC 2580  
 DB 7515 CCGGATCTGAGAGCGGAGCTTGTGTACCTGAGTGAAGAGCGGAGAGAGAGTGTGTC 7574  
 QY 2581 TGTGTCTGATGTCTTCAACATGAGAGAGCGCTCTGATCAAGCATGCGTGGAGAGAA 2640  
 DB 7575 TGTGTCTGATGTCTTCAACATGAGAGAGCGCTCTGATCAAGCATGCGTGGAGAGAA 7634  
 QY 2641 AGCAAGCTGCCATCAACGCGTGAAGCACTCTTGTCTGCTGATCAACATGCTTAT 2700  
 DB 7635 AGCAAGCTGCCATCAACGCGTGAAGCACTCTTGTCTGCTGATCAACATGCTTAT 7694  
 QY 2701 GCTACACATCCCGGAGCGGAGAGCGGAGAGAGAGTCACTTTGACATGAGAGAA 2760  
 DB 7695 GCTACACATCCCGGAGCGGAGAGAGCGGAGAGAGAGTCACTTTGACATGAGAGAA 7754  
 QY 2761 ATCTGAGAGATCACTACAGAGAGTGTCTCAAGAGATGAGAGAGAGAGTGTCTCA 2820  
 DB 7755 ATCTGAGAGATCACTACAGAGAGTGTCTCAAGAGATGAGAGAGAGAGTGTCTCA 7814  
 QY 2821 AAGGCTAAGCTTCTATGATGAGAGAGAGTGTCAAGCTGACGCCGCCCACTTCCGCCAAA 2880  
 DB 7815 AAGGCTAAGCTTCTATGATGAGAGAGAGTGTCAAGCTGACGCCGCCCACTTCCGCCAAA 7874  
 QY 2881 TCTAAATTTGGCTATGAGAGAGAGAGTGTCAAGCTGACGCCGCCCACTTCCGCCAAA 2940  
 DB 7875 TCTAAATTTGGCTATGAGAGAGAGAGTGTCAAGCTGACGCCGCCCACTTCCGCCAAA 7934  
 QY 2941 ATCCGCTCCGCTGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
 DB 7935 ATCCGCTCCGCTGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7994  
 QY 3001 ATGGCAAAAATGAGGTTTCTGCTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
 DB 7995 ATGGCAAAAATGAGGTTTCTGCTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8054

QY 3061 CTTATGCTGTTCCAGACTTGGGAGTCCGTGTGTGAGAGAGAGAGAGAGAGAGAGAG 3120  
 DB 8055 CTTATGCTGTTCCAGACTTGGGAGTCCGTGTGTGAGAGAGAGAGAGAGAGAGAGAG 8114  
 QY 3121 GTCTCAACCTCTCCCTAGAGCTGTGATGAGTCTCTGATCGAGATTCAGATTTCTCTGGA 3180  
 DB 8115 GTCTCAACCTCTCCCTAGAGCTGTGATGAGTCTCTGATCGAGATTCAGATTTCTCTGGA 8174  
 QY 3181 CAGCGGCTGAGTTCCTGAGAGAGAGCTGGAATCAAGAGAGAGAGAGAGAGAGAGAG 3240  
 DB 8175 CAGCGGCTGAGTTCCTGAGAGAGAGCTGGAATCAAGAGAGAGAGAGAGAGAGAGAG 8234  
 QY 3241 TATGACACCGCTGTTTGAAGTCAAGATGAGATGAGATGAGATGAGATGAGATGAG 3300  
 DB 8235 TATGACACCGCTGTTTGAAGTCAAGATGAGATGAGATGAGATGAGATGAGATGAG 8294  
 QY 3301 ATTATCATGATGTTGATCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
 DB 8295 ATTATCATGATGTTGATCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8354  
 QY 3361 CCGCTTATATGAGGAGTCCCTGAGATTAATCAAGAGAGAGAGAGAGAGAGAGAGAG 3420  
 DB 8355 CCGCTTATATGAGGAGTCCCTGAGATTAATCAAGAGAGAGAGAGAGAGAGAGAGAG 8414  
 QY 3421 TGGCGGAG 3480  
 DB 8415 TGGCGGAG 8474  
 QY 3481 GCCTTTCAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
 DB 8475 GCCTTTCAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8534  
 QY 3541 GGCCTGTGCTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
 DB 8535 GGCCTGTGCTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8594  
 QY 3601 TTACAGAGAGCTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3660  
 DB 8595 TTACAGAGAGCTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 8654  
 QY 3661 GACCTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3720  
 DB 8655 GACCTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8714  
 QY 3721 AAAAGGATATCTTCACTCAACCTGAGACC 3749  
 DB 8715 AAAAGGATATCTTCACTCAACCTGAGACC 8743  
 RESULT 8  
 AB049090 9573 bp RNA linear VRL 22-AUG-2002  
 LOCUS Hepatitis C virus gene for polyprotein, complete cds,  
 DEFINITION isolate:HCVT140.  
 ACCESSION AB049090  
 VERSION AB049090.1 GI:11559446  
 KEYWORDS  
 SOURCE Hepatitis C virus  
 ORGANISM Hepatitis C virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.  
 REFERENCE 1  
 AUTHORS Takahashi,K., Iwata,K., Matsumoto,M., Matsumoto,H., Nakao,K.,  
 Hatanaka,T., Ohta,Y., Kanai,K., Maruo,H., Baba,K., Hijikata,M., and  
 Mishiro,S.  
 TITLE Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients  
 with hepatocellular carcinoma: the 'progression score' revisited  
 JOURNAL Hepatol. Res. 20 (2), 161-171 (2001)  
 PUBMED 11348851  
 REFERENCE 2 (bases 1 to 9573)  
 AUTHORS Mishiro,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,

Department of Medical Sciences; 6-3-22 Higashi Oh-i, Shingawa-ku,  
Tokyo 140-8522, Japan (E-mail:shunji.mishiro@o.riken.go.jp,  
Tel:81-3-3764-8981, Fax:81-3-3764-8992)

FEATURES  
source Location/Qualifiers

1..9573  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/isolate="HCV1140"  
/specific\_host="human"  
/db\_xref="taxon:11103"  
/issue\_type="serum"  
/note="HCV-genotype: 1b"  
<1..310  
/note="partial"  
311..9343  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="BA118803.1"  
/db\_xref="GI:11559447"  
/translation="MSTNPKPQKTKRNTNRPDVYKPGGQIVGGVLLPRGRPL  
GVATRTKTSRSGRGRROPVAKRQPEGMAOPGPMPLVNGELGAGMILLSPG  
SRSPGPTDPRSRRLGVLDITLGGPADLMGTVPLVGAPLGGARALAHGRLVD  
GVNATGNIPLGCPFSIFLALLSLCTIPASVYRVRVSGYHTNDCNSIYEAAD  
MIMHTGCVPCVEDNVSRVWALTEPLAARNSIPTTIRRVLDLVGAALCSAMT  
VQDLCSGVFLVSQLFTSPRRYETQNCISYPHVSGBHMAAMMMSPPTALVQ  
SGLARIPOAVVDVAGAHGVLAGLAYSVWMAVLYLMLEFADGVHTTGGQA  
GOHALIFTSFSSGASOKIOLINTGSMHNRNALNCNSLHFLFALFPAKFNAS  
GGERMASCRPIDKPDGSGISHVVPNISDORPYCMHAPRGCIIPASOVGPGYC  
PFPSPIVVGTTDRFGVPTTGWGENETDVLNTRPQGNRECTMNGTPTKTCG  
PENCIGAGNNITLCTPDCRKHPEATYTKCSGSPWLPRECIYDPRYRIMHPTCTN  
TI PKRMVYGVGSHRLNLAACNMTGRSCNLEDRSELSPLLTSEMOIMPSTPL  
PALSTGLIHLHONIVDOYLYGVGSAVSVIKMEYVVLFLILADARVACIMMLL  
IAOBAELLENLVLANASVAGHGLTFVFCAMVYIKGLVPGAAVAYFGWPLL  
LILALPPRAYAMDRMAASCGTFVFGVILVILSPHYKVELARILMILQYTRAC  
LQWITPILAVRGRDAIILITCAVHPEIVETIKILAILGPMVQAQITRYPYVR  
AQLIRACMLVRKVGAGHYVQALMLKLAALGTIVDHLPLMDMAGQWRLIADVE  
PVIFSDMETKIITWADTRACGDIIIGLPPSARGRIILGPDSDGEOQWRLIADVE  
AYSQOQRLGLGCIITSLTGRDNQVGEVQVSTATOSFLATCNGVCWTYVHAGSK  
TLAAGKPIITOMYTNVDOLGVMPAPGARSLPCTCGSSDLVLRHADVIFVRBG  
DSRGLSLPRPISYLKSGSGPLCPSGHAGVIFRAAVCTRGVAKVDFPVPSMETT  
MRSPVTNDNSPPAPVQTPOTROYAHITHAPSGSGSTYPAAYAAQYVLYNVEVATL  
GFAVMSKAHGVDPNIRTVGRTITGAPITSTYVGLADGCGSGAGVILITDCBS  
TDSITSLIGITVLDQAEATAGARLVATATPPGSVVPVPHNIEVALISNTGILPFYK  
AILEAIIKGRHLIFCHSKKCDLAKSLGCVNAAVAYRGDVIPTSGDVVVA  
TDLAMGTGDFDSVIDCMTQVDFSLDPTFTJETTTVPDAYSRSORGRGSG  
RGTGVPVTPGERSGMLDSVLCEDYDCAWVELTPAETSVRLAYINTPELPQO  
DHLEFMSVETGLITHDAHFLSOTKQGNFPLVAYQATVCARAQAPPSNMKNC  
LIRLKLPHGPPLLYRLGAVNEVTLNHTYIMACMSADELVSTWLVGVLA  
ALAAVCLTGSVIVGRIILSGKPAIIPDEEVLYOEDMEBCASHLYIEQMOLE  
OFKQALGLQATKQALEAAPVESKMLLETFMAKHMNPFISGLIOYLAGSLTSGN  
PALASMAFASITSLPTTQHTLLENLIGWVAOALAPSAAPFVAGAGAAVSGSI  
GLKQVLDILLAGAGVAGVLA PKIMSGTSTEDLVNLLPILBSGLVYVCAA  
ILRRHVPREGAVOMNRLLAPASRGHNSPTHYVESDAAVLTQISLPLITQOLK  
RLHOMINEDSTPCSGSWLMDVMDICTVLDVFKTMVQSVLRPLPVPISLQCKG  
GWRGDGIQOTPCGAQIAGVHNGSMRLVGPRTCSNCHGVPFINAYITGCTCPR  
APVSRBALMRVAAEEVEITRVDFHYVTMTDNIKCPQVAPBEFTLDCVRLR  
YVACSKPLDEVTFOVGLNOYVVSQLEPEPDDVTLVLSMTLDPESHIAETKRELL  
ARSPSILASSASQLSAPLKAICTTROSPPADLIENALIRQMGNNITVESEN  
KVYILDSFEPILRAEBDEREYVAAEILIRKSKRPALPVARDYNAPLIESKEDPY  
VPPVHGCPLPTKAPPIPPRRKRYVLESTVSSALALATKTRSSSSVADST  
ATAPPDQSDGAGSVESYSSMPLEGGDDPLSDGSSQVSEASVCCVCSNS  
YTWGALITPCAABESKLPINALSNSILRHNNVYATTSASQORAKVTFDLQVLD  
DHYRDVLEKMKAKASTYKAKLSVEACKLTPHSARSKYGCAKQVRNLSANVNI  
RSYWKDLLEDTOTPIOTYMAKNVFCVQEKGRKPARLIVPDLGVRCEMALYD  
VSTTLPOAVWGAATGYOYSPGQVYFLVNAWKSRCMGPAYTRCPDSTVTSQDRIY  
EESIVQCCDLAEPARQARISLTERLYVGLTNSKGNCGYTRCRASGVITSCGNL  
TCYKLTAAACRAKALDCTMLVCGDDLVVYCESAGQEDASAKVFEATRTSPAPG  
DPOPEYDLELITSCSNVSVADHAGSKRYVLRDPTTLAAMAEVATRVNAML  
GNITVAPITMAAMILMTFISILAOELKALDOICVATYSIEBILDPOLITRHL  
GLSAGLSHSGPEINRVASCLRTKIGPPLRVWRHPRASRATILSCGGAAACGYL  
PNNAVRTKLTLPPLPAASQDLDSGWFIAIGSGGDIHLSLRAPRRPMCLILISGV  
GTYLLPNR"

3'UTR  
9344..9396

/note="upstream of poly-U stretch"  
9397..9439  
/note="poly(U) length is indistinct  
poly(U) stretch"  
9440..9573  
/note="3' X tail"

3'UTR  
3'UTR

Query Match 87.9%; Score 3297.8; DB 14; Length 9573;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 3467; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

QY	1	TGGAGAGCGCTTTTCACAGGCTTCACCCACGTGATGCCACTTCTGTCCTCAACAAAG	60
DB	4982	TGGAGAGGTCTTTCACAGGCTTCACCCACATGATGCCATTTCTTGTCCAGACTAA	5041
QY	61	CAGGACGAGACAACCTTCCCTACCTGAGGGGTACAGGCTATGTGCGCTAGGGCC	120
DB	5042	CAGGACGAGACAACCTTCCCTACCTGAGGGGTATACAGGCTACAGGCTGCGCAGGCT	5101
QY	121	CAGGCCCCCTCCATCATGAGGATCAAAATGTGAAAGTGTCTGATACGCTTAAAGCTTACT	180
DB	5102	CAGGCTCCACTCTCATCTGAGGATCAAAATGTGAAAGTGTCTATACGCTTAAAGCTTACG	5161
QY	181	CTGCGCGGCGCAACACCTTCTGTTATAGGCTGGAGCCGCTCCAAAGAGAGTCACTTC	240
DB	5162	CTGCAAGGGCGCAACACCTTGTGTATAGGCTAGGAGCCGCTCCAAAGAGAGTCACTTC	5221
QY	241	ACACACCCCATTAACCAATTCATCATGAGATGATGATGACGCGACCTGAGAGGCTGTCAG	300
DB	5222	ACACACCCCATTAACCAATTCATCATGAGATGATGATGATGATGATGATGATGATGATGATG	5281
QY	301	AGCAGCTGGAGTCTGCTGGGCGGGGCTCTTGAGAGCTCTGGCTGCTATTTGCTTGACA	360
DB	5282	AGCAGCTGGAGTCTGCTGGGCGGGGCTCTTGAGAGCTCTGGCTGCTATTTGCTTGACA	5341
QY	361	GCGAGCTGATCATTTGGTGTAGATCATCTTGTCCGGCGCGCTTATTTGTTCCCGAC	420
DB	5342	GCGAGCTGATCATTTGGTGTAGATCATCTTGTCCGGCGCGCTTATTTGTTCCCGAC	5401
QY	421	AGGGAAGTCTCTACAGAGATTCATGAGATGAGAAAGTGGCGCGGACCTCCCTTAC	480
DB	5402	AGGGAAGTCTCTACAGAGATTCATGAGATGAGAAAGTGGCGCGGACCTCCCTTAC	5461
QY	481	ATCGACAGGGAATGAGGCTCGCGGAGATTCAAGCAAAAAGCGCTCGGTTGCTGAC	540
DB	5462	ATCGACAGGGAATGAGGCTCGCGGAGATTCAAGCAAAAAGCGCTCGGTTGCTGAC	5521
QY	541	ACAGCCACCAAGCAGCGAGCGCTGCTCCCTGCTGGAGTCCAAGTGGCGAGCCCTT	600
DB	5522	ACAGCCACCAAGCAGCGAGCGCTGCTCCCTGCTGGAGTCCAAGTGGCGAGCCCTT	5581
QY	601	GAGACCTTCTGGCGGAAACATGTTGAACTTCATAGCGGGAATAAGTACTTAGAGAGG	660
DB	5582	GAGACCTTCTGGCGGAAACATGTTGAACTTCATAGCGGGAATAAGTACTTAGAGAGG	5641
QY	661	TTGTCACTCTGCTGGGAATCCCGGATGTGATCATGATGAGCTTCAAGGCTCTGTC	720
DB	5642	TTGTCACTCTGCTGGGAATCCCGGATGTGATCATGATGAGCTTCAAGGCTCTGTC	5701
QY	721	ACTAGCCGCTGTCACCAACCAATTCATCTCTGCTTAAATCTCTGGGGGATGGGTAGCC	780
DB	5702	ACTAGCCGCTGTCACCAACCAATTCATCTCTGCTTAAATCTCTGGGGGATGGGTAGCC	5761
QY	781	GCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	840
DB	5762	GCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	8521
QY	841	GCTGTTGGAGCATAGGCTTGGGAAGTGTGTTGGAACATCTTGGCGGGCTATGAGACA	900
DB	5822	GCTGTTGGAGCATAGGCTTGGGAAGTGTGTTGGAACATCTTGGCGGGCTATGAGACA	5881
QY	901	GAGTGGACAGGCGGCTGCTGCTGCTTAAAGTATGAGCGGCGAAATGCCCTTCCACCGAG	960





QY	3121	GTCTCACCTCCCTCAGGCTGTGATGAGGCTCTCTGACGATTCAGATTCTCTCGA	3180
Db	8102	GTCTCACCTCTCTCAGGCGCGATGAGGCGCTCATAGCATTCAGATCTCTCTGGA	8161
QY	3181	CAGCGGCTGAGTTCTCTGTGAAAGCCTGGAATCAAAGAGACCCCTATGGCTTTGCA	3240
Db	8162	CAGCGGCTGAGTTCTCTGTGAAAGCCTGGAATCAAAGAGATCCCTATGGCTTTGCA	8221
QY	3241	TATGACACCGCGTTTGTGCTCAACAGTACGAGATGATCCGTGTAGAGAGTCA	3300
Db	8222	TATGACACCGCGTTTGTGCTCAACAGTACGAGATGATCCGTGTAGAGAGTCA	8281
QY	3301	ATTATCAATGTTGTGACTTTGGCCCCGGAAGCAGACAGGCTTAAAGTGTCTCAGAG	3360
Db	8282	ATTATCAATGTTGTGACTTTGGCCCCGGAAGCAGACAGGCTTAAAGTGTCTCAGAG	8341
QY	3361	CGGCTTATATCGGGGGTCCCTGACTAATTCAAAAGGCAAACTGCGGCTATCCCGG	3420
Db	8342	CGGCTTATATCGGGGGTCCCTGACTAATTCAAAAGGCAAACTGCGGCTATCCCGG	8401
QY	3421	TGCGCGCGAGCGCGCTGCTGACGACTAGCTGGGTAATACCTTCACATGTTACTTGAAG	3480
Db	8402	TGCGCGCGAGCGCGCTGCTGACGACTAGCTGGGTAATACCTTCACATGTTACTTGAAG	8461
QY	3481	GCCTTGACAGCTGTGAGCTGCAAGCTCAGAGACTGACAGATGCTCGTGGGAGAC	3540
Db	8462	GCACATGCGGCTGTGAGCTGCAAGCTCAGAGACTGACAGATGCTCGTGGGAGAC	8521
QY	3541	GCGCTTGTCTTATCTGTGAGCGCGGAAACCCAGAGAGACCGCGGAGCTTACGAGTC	3600
Db	8522	GACCTTGTCTTATCTGTGAGCGCGGAAACCCAGAGAGACCGCGGAGCTTACGAGTC	8581
QY	3601	TTCACGAGGCTTATGACTAGTACTCTGCCCCCGGGGACCCGCCCAACAGAAATAC	3660
Db	8582	TTCACGAGGCTTATGACTAGTACTCTGCCCCCGGGGACCCGCCCAACAGAAATAC	8641
QY	3661	GACCTGAGTTGATTAATCATGATGCTCTCCAAATGTGCGGTCCGCGACAGATGATGAC	3720
Db	8642	GACCTGAGTTGATTAATCATGATGCTCTCCAAATGTGCGGTCCGCGACAGATGATGAC	8701
QY	3721	AAAAAGGTATATCTACTGACCCCTGACCC	3749
Db	8702	AAGCGGTATATCTACTGACCCCTGACCC	8730
RESULT 9			
LOCUS	HCVPOLYP	9379 bp	RNA linear VRL 15-JAN-1999
DEFINITION	Hepatitis C virus complete genome sequence.		
ACCESSION	AF000009		
VERSION	AF000009.1 GI:276397		
KEYWORDS	complete genome; core protein; E1 protein; E2 protein; NS2 protein; NS3 protein; NS4a protein; NS4b protein; NS5a protein; NS5b protein; p7 protein; polyprotein.		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
	viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1 (bases 1 to 9379)		
AUTHORS	Trowbridge, R. and Gowans, B.J.		
TITLE	Molecular cloning of an Australian isolate of hepatitis C virus		
JOURNAL	Arch. Virol. 143 (3), 501-511 (1998)		
MEDLINE	98232263		
PubMed	9572551		
REFERENCE	2 (bases 1 to 9379)		
AUTHORS	Trowbridge, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUL-1997) Trowbridge R., Hepatitis Unit, Sir Albert Skrzewski Virus Research Centre, Royal Children's Hospital, Herston Road, Brisbane, Queensland Q 4029, AUSTRALIA		
FEATURES	Source		
	1..9379		

misc\_RNA  
5' UTR  
CDS

mat\_peptide  
mat\_peptide  
mat\_peptide

```

/organism="Hepatitis C virus"
/virus
/mol_type="genomic RNA"
/isolate="HCV-A"
/db_xref="taxon:11103"
1..9379
/note="genomic RNA for polyprotein precursor, complete genome"
1..324
325..9357
/codon_start=1
/product="polyprotein"
/protein_id="CA03854.1"
/db_xref="GI:2764398"
/db_xref="GOA:O93016"
/db_xref="UniProt/TrEMBL:O93016"
/tranlation="MSTNPKPQKTKRNTNRNDPVKPPGGQIVGVYLLPRGRPL
GYRALRKTSSRSQPRGRORIPIKARRPERARMAOQGYPPPLVGNMGMAQVLLSPRG
SRPSWSPSDPRRRSRNLKVLDITLCGFADLNGYLPVGAPLGAARALADGVRLIED
GVNATGNLPGCSFSIFLGLSLCLTIPASAEVNRNSGVYVTDNCNASTVEEAD
MIMHVGCVPCVRVNDNSRCVVALPTTLAARASVPTTAIRRHVLLVGAATFCSAM
VGDLCGSVFLVQOLFESRPRRHETVODNCSTLYPDHITGHRAMDMMSPTALVV
SGLLIRPOLIYDMVAGAHWGVLAGLAAYMSNMKAVLIVMLLFGVSDTHHTGVA
GKDLIRFTFESLGRKQKIVLVNTHGSHNINFTALNCDSLTMHMLALFYTHSRAS
GCPERRMAGHPIDEPAGWGPITYABHSSSDRPPYCMHAPPCGIVPSEVCGPVTC
FPSPVAVGTTRHGVPTYSWGENGVLLNNTPBPQGNMGCTMMNGTGTCTCG
PCNIGVGNNTLTCPTDCFRKHPEATYVLCGSPMLIPRCWVDYPRMLHPYCTVNF
TIFKRMVYGVGHEHRLSACNMTGRGRCLEDRSSELSPLLSSTEMQVLPCSPFTL
PALSTLJLHONIVDOYLVGIGGVVUSFAIKMRYVLLFILLDARVACACLMMLL
INQAEALENVLVNAAVAGAHGILSTVFPCAMVYIKGRVPGAAVAYGVWPLLL
LILALPRAVYALDREMAACGAVFVGVLLTSLPHYKFLRLIMLDIYITRREAL
LDVMIPLPLNVRGRDAIILITCVHBEILFDITKLMLAGPMMVQAVITVYFVR
AOGILRACMLVRKRGAYVOMALMKMLGLSTVYDHLTPQDMAHGGLDIAVAVE
PVFSPDMETKILITWVGADTPAACGDIISGLPVSARBEILIGPADSLKEGMLLAPIT
AYSOOTRGLLGLITSLTGRDNQVYEGGQVSTVNTQSLATCNGVCTMGVTHVAGGSK
TLAEGKPTQMTVNDQDLYCMPAPRARSILPTCCGSDIYLITRHADVIPYRRG
DSRGLSLSPRPVSYLKSSGGLPCPSGHAVALFRAVCTRGAVARVDVPBSMETT
MSAPVETDSSPPAPETFOVHLAHPGSGSKTPVAAVAAQGVYVLNBSVAATL
GFGVMSKAHGPDPIRNTGRTITGAPPTVSTYGLADGCGSGGAYDIIICDECHS
TOSTTILGIVLDOAETAGARLVVATATPGSVTVPHNIEVLSNIGELFPGK
AIPETIKGSRHLIPCHSKKCDDELAATKISGLINAVAYRGIDVSVIPIGSDVVVA
TDLMTGPTGDDSDVDCNTCTQTYVDSLDPTFTIETTVPQDANBSORGRTRCG
RRGILFVTPGRSPSGMDSVLCBCTYDGCAMVETLPAETSRRLAATNTRELPCQ
DHLEFWSVETGLTHIDAPLSQTKOAGNFPYLNAVQTVCARAQPSPSDQMKC
LILAKPTLHGPTPLVRLGAVONBEVLTPIRTYIMACGSADELVVTSVMVAVGLA
AAVXCTLGTGSVVIYGRILISGRPAIIPREVLYOBFDEMEBCASHLYIEQMOLAB
OKOKALGLIQTATQAEAVAPVVSCKMALAFAPAKHMNFIISGTOYLAGISTLGN
LAIASIMAFATGITSPTLTOHTLIPNIILOGWATAOLAPPSAASAPGAGIAGAANSI
GLKVLVLDILAGAGVAGALVAFKVSSEMSTEDLVNLPALIPSGALVGVVCA
ILRHVGPBEGAVOMNRLIAPASRGNHVSPTHYVESDAARVTOILSLITLOLK
RLHONINEDCSTPSCGSMRLDWDIMCTVLPDKTLOSLPLRGLPIFYSGORGYK
GVMRGQIMOTTCPCGAOLITGVKNGSNMIVGKTSNMTHGFTFINATGTGCTSP
AVYRSPALRVAAEVEYTRVGDHRYVNTGCTMDVNCPCOVAPPEFTEVDGRLHR
YAPACKPLRLDEYTFOVGNQVYVSQLECEBPEDVYVYVTSMLTDSHITAEAKRL
DKGSPSLASSSASQSLASPLACTTRHDSPPADILIANLRLQMGNTIKVESBY
VPPVHVGCLPPTKAPPIPPPKRKTIVLTESVSALAEILATKTESSESSAVSGT
ATAPBPDPDNDPDSGDSGSMPLDEBPDPDLSGMSVSEAEEDVYVSCMS
YTMGALLTPCAABESKLPINLSNGLIHHNMVYATRSASOROKKYTPRLTOYLD
DHTRDLKEMKAKASTVKALLSVEBACKLTPHRSKRSFGYGAQDVRNLSKAVNHI
RSWYKDLKEDTEPIPTITMARNEVECVQPEKGRAPALIVPDLGVVCEMALYD
VSTLTPQAVMGSPGYOYSPQVEFLVNAVMSKCPMPGAFADVTRCFDVTJESDTRV
BESIYOCQALBPBAROAKISLTERLYIGPLTNSKONGRCYRCSGLTSCGTL
TCYKASACRAKIQDCTMIVOGDVLIVICISAGQEDASARVTEAMTRHSPVAGP
DPQPPYDIELITSCSNVSVADHAGSKVYVITRPTPLAARAAETARHPVNSML
GNITMYAFTLMAAMILMTHFSFLLAQEOLKRLDCQITGACSTIEPLDIPOLIOKLH
GLSAPLSHSYSGEIRIVASCLRKGVPLRARHRARVRAVLLISQGGGAATCGVYL
FMAVSKTKLITPIPAASOLDLSNMFVAGSGDIIHYSIRAPRFRMCMCLLSVGV
GIYILNRR"
325..897
/product="core protein"
898..1473
/product="E1 protein"

```

mat\_peptide 1474..2562  
 /product="E2 protein"  
 mat\_peptide 2563..2751  
 /product="p7 protein"  
 mat\_peptide 2752..3402  
 /product="NS2 protein"  
 mat\_peptide 3403..5295  
 /product="NS3 protein"  
 mat\_peptide 5296..5457  
 /product="NS4 protein"  
 mat\_peptide 5458..6240  
 /product="NS4b protein"  
 mat\_peptide 6241..7581  
 /product="NS5a protein"  
 mat\_peptide 7582..9354  
 /product="NS5b protein"  
 3'UTR  
 9358..9379

Query Match 87.9%; Score 3294.6; DB 14; Length 9379;  
 Best Local Similarity 92.4%; Pred. No. 0;  
 Matches 3465; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

QY 1 TGGAGGCGCTCTTACAGGCTTCACCCAGTGGATGCCACTTCTGCCAAACAAAG 60  
 DB 4996 TGGAGAGCGCTTTCACAGGCTTCACCCATAGAGCCCACTTCTGCCAGACCAA 5055

QY 61 CAGGACGAGAACAACTTCCCTTACCTGAGCGCTACCAAGCTACTGTGTGCGCTAGGCC 120  
 DB 5056 CAAGCGGAGAGAACTTCCCTTACCTGACAGCGTACCAAGCTACTGTGTGCGAGGCT 5115

QY 121 CAGGCCCCACCTCATCATGAGATCAAAATGGAAGTGTCTATAGGCTTAAAGCTACT 180  
 DB 5116 CAGGCTTCACCTCATCTGTGAGATCAAAATGGAAGTGTCTATAGGCTTAAAGCTACT 5175

QY 181 CTGCGGGGGCAACACCTTGTCTATAGGCTGAGAGCCCTCCAAACGAGCTCACCTC 240  
 DB 5176 CTGACGGGGCAACACCTTGTCTATAGGCTGAGAGCCCTCCAAACGAGGCTGTCTT 5235

QY 241 ACACACCCCATTAACCAATTCATCATGAGATGATGATGATGATGATGATGATGATG 300  
 DB 5236 ACACACCCCATTAACCAATTCATCATGAGATGATGATGATGATGATGATGATGATG 5295

QY 301 AGCAGCTGGGTGTGATGAGGCGGGGCTCTTGAAGCTGTGCTGGTATGCTTGAACA 360  
 DB 5296 AGCAGCTGGGTGTGATGAGGCGGGGCTCTTGAAGCTGTGCTGGTATGCTTGAACA 5355

QY 361 GGCAGCGTGTCTATGATGAGATCATCTTGTCCGGGCGGCGCTATTGTTCCGAC 420  
 DB 5356 GGCAGCGTGTCTATGATGAGATCATCTTGTCCGGGCGGCGCTATTGTTCCGAC 5415

QY 421 AGGGAAGTCTCTTACAGAGATTCATGATGATGATGATGATGATGATGATGATGATG 480  
 DB 5416 AGGGAAGTCTCTTACAGAGATTCATGATGATGATGATGATGATGATGATGATGATG 5475

QY 481 ATCGACAGGGAATGAGCTCGCGGAGCTTGAAGAAACCGCTCGGGTGTCTGACG 540  
 DB 5476 ATCGACAGGGAATGAGCTCGCGGAGCTTGAAGAAACCGCTCGGGTGTCTGACG 5535

QY 541 ACAGCCACCAAGCAACGAGGCGGCTGTCTCCGCTGTGAGATTCAGAGGCGGCTT 600  
 DB 5536 ACAGCCACCAAGCAACGAGGCGGCTGTCTCCGCTGTGAGATTCAGAGGCGGCTT 5595

QY 601 GAGACTTCTGTGGGCAACATGATGAACTTATCATGAGCGGATATCATGATGATGATG 660  
 DB 5596 GAGACTTCTGTGGGCAACATGATGAACTTATCATGAGCGGATATCATGATGATGATG 5655

QY 661 TTGTCACTGTGCTGAGGAATCCGCGATGATGATGATGATGATGATGATGATGATG 720  
 DB 5656 TTGTCACTGTGCTGAGGAATCCGCGATGATGATGATGATGATGATGATGATGATG 7315

QY 721 ACTAGCCGCTCAACCAATTAACCTCTCTCTTAAATCTCTGGGGGATGGATGAC 780

DB 5716 ACCAGCCGCTCAACCAACCAATACCTCTCTGTTTAACTTGGGGGATGGGTGCT 5775

QY 781 GCCCACTGCTCCCCCAAGTGTCTTACGCTTTGTATAGCGCCGCAATTCGTGTCG 840

DB 5776 GCCCACTGCTCCCCCAAGTGTCTTACGCTTTGTATAGCGCCGCAATTCGTGTCG 5835

QY 841 GCTGTGGCAGCATTAAGGCTTGGGAAGTCTTGTGACATCTTGGCGGCTATGAGCA 900

DB 5836 GCGGTGGCAGCATTAAGGCTTGGGAAGTCTTGTGACATCTTGGCGGCTATGAGCA 5895

QY 901 GAGTGGCAGGCGGCTGTGCTTGAAGTATGAGCGGAAATGCTTCCACCGAG 960

DB 5896 GAGTGGCAGGCGGCTGTGCTTGAAGTATGAGCGGAAATGCTTCCACCGAG 5955

QY 961 GACCTGGTTAACTTAATCTCTGCAATCTCTCTGTCGCTTGTGTCGCTGTCGTCGTC 1020

DB 5956 GACCTGGTTAACTTAATCTCTGCAATCTCTCTGTCGCTTGTGTCGCTGTCGTCGTC 6015

QY 1021 TGGCAGCGATATGCGTGGGACGATGGGTCCAGGGGAGGGGCTGTGAGTGAAC 1080

DB 6016 TGGCAGCGATATGCGTGGGACGATGGGTCCAGGGGAGGGGCTGTGAGTGAAC 6075

QY 1081 CGGCTAATAGCGTTCGCTGGGGGTAACATGTTTCCCAAGCAATATGTCAGAG 1140

DB 6076 CGGCTAATAGCGTTCGCTGGGGGTAACATGTTTCCCAAGCAATATGTCAGAG 6135

QY 1141 AGGAGCGGCGACAGTGTCACTAGATCTCTCCGACTTATCATCCCACTGTG 1200

DB 6136 AGGAGCGGCGACAGTGTCACTAGATCTCTCCGACTTATCATCCCACTGTG 6195

QY 1201 AAGAGCTCCACAGTGGATTAACAGAGAGCTGCTCCAGCGCCCTGCTGGTGGCTA 1260

DB 6196 AAGAGCTCCACAGTGGATTAACAGAGAGCTGCTCCAGCGCCCTGCTGGTGGCTA 6255

QY 1261 AGGAGTGTGGGACTGATATGACAGATTTTGGCTGACTTCAAGACTTGGCTCAGTCC 1320

DB 6256 AGGAGTGTGGGACTGATATGACAGATTTTGGCTGACTTCAAGACTTGGCTCAGTCC 6315

QY 1321 AAGCTCTGCGGATTAACGGGAGTCCCTTTTTCATGCCAAGTGGGTCAAGGGG 1380

DB 6316 AAGCTCTGCGGATTAACGGGAGTCCCTTTTTCATGCCAAGTGGGTCAAGGGG 6375

QY 1381 GTCGGCGGGGAGACGGCATCATGACAGCAACCTGCTCATGTGGAGACAGATCACCG 1440

DB 6376 GTCGGCGGGGAGACGGCATCATGACAGCAACCTGCTCATGTGGAGACAGATCACCG 6435

QY 1441 CATGTCAAAAACGGTTCATAGAGATGTTGGGCTTAAGACTGTATTAATGTGACAT 1500

DB 6436 CATGTCAAAAACGGTTCATAGAGATGTTGGGCTTAAGACTGTATTAATGTGACAT 6495

QY 1501 GGAACATTCCTCATCAAGCATACAGAGGCGCTGCAAGCTCTCCAGCGCCAAAC 1560

DB 6496 GGAACATTCCTCATCAAGCATACAGAGGCGCTGCAAGCTCTCCAGCGCCAAAC 6555

QY 1561 TATTCAAGGGCGTGTGGGGGTGCTGCTGAGGAATGATGAGGTTACGCGGGGTGG 1620

DB 6556 TATTCAAGGGCGTGTGGGGGTGCTGCTGAGGAATGATGAGGTTACGCGGGGTGG 6615

QY 1621 GATTTCCATAGTGAAGCATGACCATGACAAAGTAAATGACCGGTGCGAGTTTCCA 1680

DB 6616 GATTTCCATAGTGAAGCATGACCATGACAAAGTAAATGACCGGTGCGAGTTTCCA 6675

QY 1681 GCGCCGGAATTTTCAAGAGTGAATGGGGGTGCGGCTGCAAGTACCTCGCGGTGC 1740

DB 6676 GCGCCGGAATTTTCAAGAGTGAATGGGGGTGCGGCTGCAAGTACCTCGCGGTGC 6735

QY 1741 AAACCTCTCTTACGGAGAGGTCATTCAGGTGGGCTCAACCAATATCTGTTGG 1800

DB 6736 AAACCTCTCTTACGGAGAGGTCATTCAGGTGGGCTCAACCAATATCTGTTGG 6795

QY 1801 TCGAGCTCCATGACAGCGCAACCGATGATGAGAGTGTCACTTCCATGCTCACGAC 1860

DB 6796 TCGAGCTCCATGACAGCGCAACCGATGATGAGAGTGTCACTTCCATGCTCACGAC 6855

QY	1861	CCCTCCCAATTCACAGCAGAGACCGCTTAAGGCGAGCGTGGCAGAGGGGTCTCCCTCC	1920
Db	6856	CCCTCCCAATTCACAGCAGAGACCGCTTAAGGCTGAAGCTGAACAGAGGGGTCTCCCTCC	6915
QY	1921	TTGGCAGACTCTTCAGCTTAGCCAGTTGTCGTGCGCTTCTCGAAGCGCACATTAATTACC	1980
Db	6916	TTGGCAGACTCTTCAGCTTAGCCAGTTGTCGTGCGCTTCTCGAAGCGCACATTAATTACC	6975
QY	1981	CAAAATGACTTCCCGAGAGCTGACTCTCATTCGAGGCGCACTTCTCTGTGCGGCATGGAATG	2040
Db	6976	CGTCACGACTCCCGAGAGCTGACTCTCATTCGAGGCGCACTTCTCTGTGCGGCATGGAATG	7035
QY	2041	GGCGGAGGACATTAACCGCGTGGAGTGCAGAGAAACAAGTAGATATCTTGACTCTTTGCAC	2100
Db	7036	GGCGGAAACATCAACCGCGTGGAGTGTGAATCTGAATAACAAGTAGATATCTTGACTCTTTGCAC	7095
QY	2101	CCGCTCCGACGGAGAGATGAGCGGGAAGTGTCCGTCCCGCGGAGATCTTCGCGAAA	2160
Db	7096	CCGCTTCAGCGGAGAGATGAGGAGGAATATCCGTCCGCGGAGATCTTCGCGAAA	7155
QY	2161	TCCAGGAATTCGCCACGACGATGCGCGCATGGGAGCGCCGGAATTACAACCTCCGCGTG	2220
Db	7156	TCCAGGAATTCGCCCGAGCGATGCCCATATGGGCAACGCGGATTAACAACCCCTCACTG	7215
QY	2221	CTGAGTCTGTGAAGAGCCCGGACTAGCTCCCTCCAGTGTGATCAATGGGTGCCACTGCCA	2280
Db	7216	CTAGAGTCTGTGAAGATCCCGGACTAGCTCCCTCCGAGTGTGACAGGGGTGCCATTACCA	7275
QY	2281	CCTACTAAGACCCCTCTTATACACTCTCAACGAGAGAAAGACAGTTGTTCTGAACGAA	2340
Db	7276	CCTACCAAGGCCCCCTTCAATACACTCTCAACGAGAAAGAGACGGTGTCTTACACAGAG	7335
QY	2341	TCCACCGTGTCTTCTGCGCTGGCGAGCTTGCCACAAAGGCTTTTCGTAGCTCCGAACCG	2400
Db	7336	TCCACCGTGTCTTCTGCGCTTGGCGGAGCTTGTCAAAGACCTTTCGAGCTCCGAATCG	7395
QY	2401	TGCGCCGTGCAGACGCGGCAACGCGCCCTCTTGAACAACTCTCCGACGACGGCGGA	2460
Db	7396	TGCGCCGTGCAGACGCGGCAACGCGCCCTCTTGAACAACTCTCCGACGACGGCGGA	7455
QY	2461	GCAGAGTCTGACGTTGAGTCGATATCTTCACATGACCCCTTGAAGGGGAGACCGGAGGAC	2520
Db	7456	ACAGAGTCCGACGTTGAATCGTGTCTCTTCATGACCCCTTGAAGGGGAGACCGGAGGAT	7515
QY	2521	CCCGATCTTCAGCAGCGGCTCTTGATCTACCGTAGTAGAAGCCGATGAAGACGTCTGC	2580
Db	7516	CCCGATCTTCAGCAGCGGCTCTTGATCTACCGTAGTAGAAGCCGATGAAGACGTCTGC	7575
QY	2581	TGCTGTCTGATTTCTCTTACATAGACAGGCGTCTGATCAACGCCATGGGCTGGCGAGGAA	2640
Db	7576	TGCTGTCTGATTTCTCTTACATAGACAGGCGTCTGATCAACCAATGCGCGCGGAGAAA	7635
QY	2641	AGCAAGCTGCGCATCAACGCGTTGAGCAACTTTTGTGCTGCTACCAACAATGGTCTAC	2700
Db	7636	AGCAAGCTGCGCATCAAGCGTTGAGCAACTTTTGTGCTGCTACCAACAATGGTCTAC	7695
QY	2701	GCTACCACTCCCGCAGCGCAAGCCAGGGGCGAAGAAAGTCACTTTGACAGACTGCA	2760
Db	7696	GCCACCACTCCCGCAGCGCAAGCCAGGGGCGAAGAAAGTCACTTTGACAGACTGCA	7755
QY	2761	ATCTCTGAGAGATCACTACAGGACGTCCTCAAGAGATGAAGAGCGCAAGGCGCCACAGTT	2820
Db	7756	ATCTCTGAGAGATCACTACCGGAGCTGTCTCAAGAGATGAAGAGCGCAAGGCGCCACAGTT	7815
QY	2821	AAAGCTAAGCTTTCTACATAGAGAAAGCTTCAAGCTTGAACGCCCTCCACATTCGGCCAA	2880
Db	7816	AAAGCTAAGCTTTCTATCCGTGAAGAAAGCTCTCAAGCTGACACCCCTCCATTCGGCCAG	7875
QY	2881	TCTTAATTTGGCTATGGGCGAAGACGTTCGCGAACCTATCCAGAAAGGCACTTAACAC	2940
Db	7876	TCTTAATTTGGCTATGGGCGAAGACGTTCGCGAACCTATCCAGAAAGGCGCTTAACAC	7935

QY	2941	ATCCGCTCCGCTGGGAGGACTTGTGGAAAGACCTGAAAACCAATTTGACACACATC	3000
Db	7936	ATCCGCTCCGCTGGGAGGACTTGTGGAAAGACCTGAAAACCAATTTGACTACATC	7995
QY	3001	ATGGCAAAAATGAGGTTTCTGCGTCCAACTAGAGAGAGAGGCGCAAGCCAGCTCGC	3060
Db	7996	ATGGCAAAAATGAGGTTTCTGCGTCCAACTAGAGAGAGAGGCGCAAGCCAGCTCGC	7996
QY	3061	CTTATCGTGTCCAGACTTGGGGGTCCGTGTGTGCGAAGAAATGGCCCTTATGACGTG	3120
Db	8056	CTTATCGTGTCCAGACTTGGGGGTCCGTGTGTGCGAAGAAATGGCCCTTATGACGTG	8115
QY	3121	GTCGCCAACCCCTCCGACGGCTGTATGGGGCTCCCTGTACGGATTCGATATTCCTCCGGA	3180
Db	8116	GTCGCCAACCCCTCCGACGGCTGTATGGGGCTCCCTGTACGGATTCGATATTCCTCCGGA	8175
QY	3181	CAGCGGGTGCAGTTCTCTGGTGAACGCTCGAATAACAAGAGAACCCCTATGGCTTTGCA	3240
Db	8176	CAGCGGGTGCAGTTCTCTGGTGAATAAGCTCGAATAACAAGAGATGCTCTATGGCTTGCA	8235
QY	3241	TATGACACCCGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA	3300
Db	8236	TATGACACCCGCTGTTTGACTCAACAGTCACTGAGAGTGCATCCGTGTAGAGAGTCA	8295
QY	3301	ATTATCAATGTGTGACTTGGCCCCCGAAGCCAGACAGAGCATTAAGGTGCTCAGAG	3360
Db	8296	ATTATCAATGTGTGACTTGGCCCCCGAAGCCAGACAGAGCATTAAGGTGCTCAGAG	8355
QY	3361	CGGCTTTATTCGGGGGGTCCCTCTGACTTAATCAAAAAGGCGAGAACTGCGCTATCGCCG	3420
Db	8356	CGGCTTTATTCGGGGGGTCCCTCTGACTTAATCAAAAAGGCGAGAACTGCGCTATCGCCG	8415
QY	3421	TGCGCGCGAGCGCGGTGCTGACGACTAGCTGCGGTAAATACCTCACAATTTACTTGAAG	3480
Db	8416	TGCGCGCGAGCGCGGTGCTGACGACTAGCTGCGGTAAATACCTCACAATTTACTTGAAG	8475
QY	3481	GCTCTGACGCTGTTCGAGCTGCAGAAAGTCCAGAGCTGCACAGATGCTGTGTGGGAGAC	3540
Db	8476	GCTCTGAGGCGCTGTTCGAGCTGCAGAAAGTCCAGAGCTGCACAGATGCTGTGTGGGAGAC	8535
QY	3541	GAGCCTTTCGTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGGCGAGCTTACGAGTC	3600
Db	8536	GAGCCTTTCGTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGGCGAGCTTACGAGTC	8595
QY	3601	TTCAAGAGAGCTATGACTAGGTACTCTGCGCCCTCCCGGGAGCCCGCCCAACAGAAATAC	3660
Db	8596	TTCAAGAGAGCTATGACTAGGTACTCTGCGCCCTCCCGGGAGCCCGCCCAACAGAAATAC	8655
QY	3661	GACCTGAGGTTGATTAACATCATGTCTCTCCCAATGTTCGGTCGGGACAGATGACTGGC	3720
Db	8656	GACCTGAGGTTGATTAACATCATGTCTCTCCCAATGTGTGGTCGGGACAGATGACTGGC	8715
QY	3721	AAAAAGGTATTACTACTCAGCCCGTGAACC 3749	
Db	8716	AAAGAGATTACTACTCAGCCCGTGAACC 8744	
RESULT 10			
AF207753			
LOCUS	AF207753	9374 bp	RNA linear VRL 27-Apr-2000
DEFINITION	Hepatitis C virus strain MD12 complete genome.		
VERSION	AF207753		
KEYWORDS	AF207753.1 GI:7650223		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	Hepatitis C virus		
AUTHORS	Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and Sato, C.		
TITLE	Characteristics of hepatitis C viral genome associated with disease progression		

JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 9374)  
 AUTHORS Nagayama, K., Kurotsaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and Sato, C.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan

FEATURES  
 source Location/Qualifiers  
 1. .9374  
 /organism="Hepatitis C virus"  
 /mol\_type="genomic RNA"  
 /strain="MD12"  
 /db\_xref="taxon:11103"  
 /note="isolated from a patient with persistently-normal ALT values (ASC-2)"  
 ALT values (ASC-2)"  
 330. .9362  
 /codon\_start=1  
 /product="polyprotein"  
 /protein\_id="AF65943.1"  
 /db\_xref="GI:7650224"  
 /translation="MSTNPKPQKTKRNTNRPPDYKFGGCGIIVGGVILPRGRRLGVATRTSERQSPQRROPPIPRVQPEGTMAQPEPMPVLYNENGGAGMILLSPRGSRBSMGTPDPRRRSRNIGKVIDLTGCFADLMGYIPVGAFLGAAAYLAHGVRLDGVNATGNLPGCSFSIFLLALLSLCTTPASAYEVNNSGMYHTNCCNSIYEAADMIHARGCVCVENNSRCWVALPTLARNASVPTTIRRHVDLIVGAARCSAMYVDLCCSVPLISQLFTPSRRHETVDDCNSITPGHVSGRNADMMNMSPTALIVSOLRLIQAIVDVAGAHMGLAYISAGMMAVLIWMLFAGVDGTRISGGTQGHHTSGVLSLFSSGSSQNIQLINTGSMHNTALNCNSLNTFLAALIMLQYLIRAAHLGATVPLNVRGGRDAIIFLCATHEPILFDITKLALIGLMLVQAGITRVPYVRAAGLIRACMLVRKAGVYOMAEMKLAALTGVYVDHLITPLRGMATGTRDLAIVAPVFSDMETKITIHWAGDTPAACGDIILGLPISARGRBILIGPDSLEGQWRLAIIITAVSQTRGLGICITISLTGRDKNOVEGVVVSATQTELATVNEVCWTFVHGAASKTLAAGKPIQTOMTNDODLVQWAPPGASMTPTCGSDLLVTRHADVLEVRBGDGRGSLSPRPVYLYKSSGSGPLCGSHVGIIPRAVCGAKAVDEPVSMTTMSPPVTNDSNPRAVPOTROVAHHAPTGSGKSTKPAVAAAGYKVLVNPVAAITLGRGAVMSKAVGTDPNIRTGRTITTGAPITSTYVGEFLADGSGSGAYDIIICDECHS785TSTLTIGTGLVDQAEFTAGARLVLAATATPPGSAVTPHPIEVALSNTGEIPFYGKALPIETIKGRHLIFCHSKKKCCDELAASKALGINAVAYRGLDVAIVPTSGVVVATDALTGTGDPDSVIDCNTVQTVDFSLDPTFITIDTTPVADVAISRTORGRTRSGIYRFTVPEGRSPGMDPSVLCEDYAGCAWELTAPETSRVLAAYLNTPELPHYOCDLHEFEGVFTGLTRIDAHPLSOTKQAGNPPYVAVQATVCRAGAAPPSPMOOMKCLTRIKPTLGGPPLIYRIGAVONENLTILHITKYIMTCHSADLEVSTWVLYGVIALAALAYCLTTGSAVIVGRITLSGRPAIIPREVLIRFEDNEBCASQULPIBOQMOLAEQERQKALGLLQTRTKQAEAAAPVESKMRLEAFWKHMNFISGIOYLAGSTLLEKNPAIASLMAFTASVSTPLTOSTLTFNLIIGWVAOLAAPSASAPFAGAGVAGVCSIGKSLVLDIILAGAGVAGALVAFKVMGSEMPSTEDLVNLPILPGLAVVONAAIILRBHVGPGEAVOMNRLIAFASRGHNSPTHYVESDLSALAVTOIILNLTITOLLRLHOWINEDCSTPGCGSMRLDWMWDTCTYLSDKPTWLSQSLPRLRGLVPLSQRKYGWRGDIQITTCGCAQITGHVKNSSMKIVGPKTSNTHGTFPINAITGCTSPAPYASRLMRVAEEVEVTRVGDVHYVGMETDNCPCQVPAPEFFEDVAKRLRYAPACKPDLREBVTPOVGNQYVLSGLCEPEPDVAIVLCTDSHLEIAETAKRLHARGSPSLASSSSASQASPLSKATCTTHDSPADLIEANLIROMGQNTIIVSEENKYILDSFPDLRABEDERISVPAELIRKRPKPPALPIIWARDVNPPLLESKMDIVPPVYHGCPLPTKPGPPLPPRKRITVLTSTVSALAEALATKGTSSGSSASDGTATAPPTQDSDGSDVESISMPLEBGPDPDLSGDSWSTVSEBAGDVCCMSYTMGTALITPCAEBESKLPINALSNLRLHHNNVATITRSACORRKAFTPDLQVLDHYRDLVKEMKAKASTVAKALVBEACKLTTPHSAKKEGYCAKVRNLSKAVNLRSEVMKDLKEDTEPTINTTMAKNEVCEVPEKGRPARLIIVPDVGRCEMAALYDVSTLPOAVGSSGYOYSGORVEFLNPMWKSCKCMGVSYTRCFDSVSTNDRVRESIYQACDLAPBAROVIRISLTERLVVGLPLTMSKCONGQYRCRCISYLTSCMQLTCTLKASACRAKLODCTMLVNGDVLVITCESAGTQEDASIRVTEAMTRISAPGDPOPEYDELITSSCSNVSVAHDAGKRVLYLTPPTPLAANAEIYLAHPTBENLGNINIAVPTLMAAMIIMTHFSLIAOBEKALDCOIGACASIEPLDLPJLIEHILGSAFSLHSGPBEINRVASCRKLCVPLRVRHRAARSRAILLGCGRAACGKXYLFMAVTRTKLKTPIPAASRLDLSGFMVAGSGDDIYHLSBRAPRMFMCLLLLSGV

ORIGIN  
 GYLLENR"  
 Query Match 87.8%; Score 3293; DB 14; Length 9374;  
 Best Local Similarity 92.4%; Pred. No. 0;  
 Matches 3464; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

1 TGGAGAGGCGCTTCAAGGCGCTCACCGCAGTGAAGCCACTTCTGTCGCAACAAAG 60  
 DB TGGAGAGGCGCTTCAAGGCGCTCACCGCAGTGAAGCCACTTCTGTCGCAACAAAG 5060

61 CAGGACGAGACAACTTCCCTCACTGATGCGCTACAGGCTACTGTGCGCTAAGGCC 120  
 DB CAGGACGAGACAACTTCCCTCACTGATGCGCTACAGGCTACTGTGCGCTAAGGCC 5120

121 CAGGCGCCACCTTCATCATGAGGATCAATGTGGAAGTGTCTCATACGCTAAAGCTACT 180  
 DB CAGGCTCCACCTTCATCATGAGGATCAATGTGGAAGTGTCTCATACGCTAAAGCTACT 5180

181 CTGGGCGGCGCAACCGCTTGTGATGAGGCTGGGAGCGGCTCCAAACGAGGTCACCGTC 240  
 DB TTGACAGGCGCAACCGCTTGTGATGAGGCTGGGAGCGGCTCCAAACGAGGTCACCGTC 5240

241 ACAACCCCATTAACAAATTCATCATGAGATGATGATGACCCGACCTGAGAGTGTACAG 300  
 DB ACAACCCCATTAACAAATTCATCATGAGATGATGATGACCCGACCTGAGAGTGTACAG 5300

301 AGCACTGGAGTGTGTGGGCGGAGGTCCTTGACAGCTTGTGGCTGTATGCTTGAACACA 360  
 DB AGCACTGGAGTGTGTGGGCGGAGGTCCTTGACAGCTTGTGGCTGTATGCTTGAACACA 5360

361 GAGAGGCTGTCAATTTGGGTAGATCATCTGTCCGGCGCGCGGCTATTTGTCGCGAC 420  
 DB GAGAGGCTGTCAATTTGGGTAGATCATCTGTCCGGCGCGCGGCTATTTGTCGCGAC 5420

5361 GAGAGGCTGTCAATTTGGGTAGATCATCTGTCCGGCGCGCGGCTATTTGTCGCGAC 5420

421 AGGAAAGTCTCTTACAGAGAGTTCATGAGATGGAAGAGTGCAGCTCCCTTAC 480  
 DB AGGAAAGTCTCTTACAGAGAGTTCATGAGATGGAAGAGTGCAGCTCCCTTAC 5480

5421 AGGAAAGTCTCTTACAGAGAGTTCATGAGATGGAAGAGTGCAGCTCCCTTAC 5480

5481 ATGGAACAGGGAATGACGCTCCGCGAGCAATTCAAACAGAGCGCTCGGATGTCGCA 5540

541 ACAGCCACCAAGCAACCGAGCGCGCTGCTCCCGTGTGAGAGTCCAAAGTGGCGAGCCTT 600  
 DB ACAGCCACCAAGCAACCGAGCGCGCTGCTCCCGTGTGAGAGTCCAAAGTGGCGAGCCTT 5541

5541 ACAGCCACCAAGCAACCGAGCGCGCTGCTCCCGTGTGAGAGTCCAAAGTGGCGAGCCTT 5541

601 GAGACCTTCTGGGCGAAACACATGTGGAACCTTCATCAGCGGAGTACAGTACTTAGCAGGC 660  
 DB GAGACCTTCTGGGCGAAACACATGTGGAACCTTCATCAGCGGAGTACAGTACTTAGCAGGC 5601

5601 GAGACCTTCTGGGCGAAACACATGTGGAACCTTCATCAGCGGAGTACAGTACTTAGCAGGC 5660

5661 TTGTCACTCTGCTGGGAACCCCGCAATGACATCAGTATGAGCATTCACAGGCTCTGTC 5720

721 ACTAGCCGCTCACACCCCAATCTACCTCTCTCTTAATCATCTGGGGGAGTGGTACGC 780  
 DB ACTAGCCGCTCACACCCCAATCTACCTCTCTCTTAATCATCTGGGGGAGTGGTACGC 5721

5721 ACCAGCCGCTCACACCCCAATCTACCTCTCTCTTAATCATCTGGGGGAGTGGTACGC 5780

781 GCCCACTGCTGCTCCCGCAAGTGTGCTTACGCTTCTGATGAGCGCCGAGATTTGCTGTGC 840  
 DB GCCCACTGCTGCTCCCGCAAGTGTGCTTACGCTTCTGATGAGCGCCGAGATTTGCTGTGC 5781

5781 GCCCACTGCTGCTCCCGCAAGTGTGCTTACGCTTCTGATGAGCGCCGAGATTTGCTGTGC 5840

841 GCTGTGGACGATAGGCTTGGGAAGTGTCTTGTGACATCTTGGCGGCTTATGAGACA 900  
 DB GCTGTGGACGATAGGCTTGGGAAGTGTCTTGTGACATCTTGGCGGCTTATGAGACA 5841

5841 GCTGTGGACGATAGGCTTGGGAAGTGTCTTGTGACATCTTGGCGGCTTATGAGACA 900

901 GAGAGTGGACGAGCGGCTGCTGCTTAAAGTGCATGAGAGGCGGAAATGCGCTCCACCGAG 960  
 DB GAGAGTGGACGAGCGGCTGCTGCTTAAAGTGCATGAGAGGCGGAAATGCGCTCCACCGAG 5901

5901 GAGAGTGGACGAGCGGCTGCTGCTTAAAGTGCATGAGAGGCGGAAATGCGCTCCACCGAG 960

961 GACCTGTTAACTTACTCTCCCTGCACTCTCTCTGCTGAGCGCTGCTGCGGAGTCGAG 1020

Db 5961 GACCTGTTAAATTACTCCCTGCCATCTCTCTCTGAGCCCTGGTGGTGGGAGTCTG 6020  
Qy 1021 TGGGAGCAGTACTGAGTGGGACGTGGGTCAGGGGAGGGGGCTGGAGTGGATGAAC 1080  
Db 6021 TGGGAGCAGTACTGAGTGGGACGTGGGTCAGGGGAGGGGGCTGGAGTGGATGAAC 6080  
Qy 1081 CGGCTGATAGCGTTCGCTCGCGGGGTAACATGTTTCCCGACGACTATGTCAGAG 1140  
Db 6081 CGGCTGATAGCGTTCGCTCGCGGGGTAACATGTTTCCCGACGACTATGTCAGAG 6140  
Qy 1141 AGGAGCGCGGACAGTGTCACTAGATCTCTCCGACTTACTATCAACCACTGGT 1200  
Db 6141 AGGAGCGCGGACAGTGTACTAGATCTCTCCGACTTACTATCACTGAGCTGTG 6200  
Qy 1201 AAGAGCTCCACAGTGGATTAAGAGGACTGCTCCAGCGCCGCTCCGGCTGTGGCTA 1260  
Db 6201 AAGAGCTCCACAGTGGATTAAGAGGACTGCTCCAGCGCACTCCGGCTGTGGCTA 6260  
Qy 1261 AAGGATGTTTGGAGTGAATGACAGTGTGGCTTGAAGACTTCAAGACTGGCTCAATCC 1320  
Db 6261 AAGGATGTTTGGAGTGAATGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6320  
Qy 1321 AAGCTCTGTGCGGATTAACGGGAGTCCCTTTTCTCATGCAAGCTGGGTCAAGGG 1380  
Db 6321 AAGCTCTGTGCGGCTGCGGGAGTCCCTTTCTCATGCAAGCTGGGTCAAGGG 6380  
Qy 1381 GTGTGGCGGGGAGACGGCATCATGACCACTGTCTCATGTGAGAGACGATCACCGGA 1440  
Db 6381 GTGTGGCGGGGAGACGGCATCATGACCACTGTCTCATGTGAGAGACGATCACCGGA 6440  
Qy 1441 CATGTCAAAAACGGTTCATAGAGATGTTGGGCTTAAAGCTGTAGTAACTATGGCAT 1500  
Db 6441 CATGTCAAAAACGGTTCATAGAGATGTTGGGCTTAAACCTGTAGCAACAGTGGCAT 6500  
Qy 1501 GGAACATTTCCCATCAACGATACCAACGAGGCTTGAACGCTTCCAGGCGCAAC 1560  
Db 6501 GGAACATTTCCCATCAACGATACCAACGAGGCTTGAACGCTTCCAGGCGCAAC 6560  
Qy 1561 TATTCAAGGGGCGTGTGGCGGGTGGCTGTGAGAGTGAAGTGAAGTGAAGTGAAG 1620  
Db 6561 TATTCAAGGGGCGTGTGGCGGGTGGCTGTGAGAGTGAAGTGAAGTGAAGTGAAG 6620  
Qy 1621 GATTTCCACTACGTGACGACATGACCATGACCACTGACCACTGACCACTGACCACT 1680  
Db 6621 GATTTCCACTACGTGACGAGGACATGACCATGACCACTGACCACTGACCACTGAC 6680  
Qy 1681 GCGCCGCAATTTCTTACAGAAAGTGAAGTGGGTGGCTGCAAGTACGCTCCGCGTGC 1740  
Db 6681 GCGCCGCAATTTCTTACAGAAAGTGAAGTGGGTGGCTGCAAGTACGCTCCGCGTGC 6740  
Qy 1741 AAACCTCTCTTACAGGAGAGTGCATCTCAGGTGGGCTCAACCACTGAGTGGG 1800  
Db 6741 AAACCTCTCTTACAGGAGAGTGCATCTCAGGTGGGCTCAACCACTGAGTGGG 6800  
Qy 1801 TCGAGCTCCCATGAGGCGGCAACCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860  
Db 6801 TCGAGCTCCCATGAGGCGGCAACCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6860  
Qy 1861 CCTTCCACATCAGAGAGAGCGGCTAAGGCGAGCTGGGCGAGGGGCTTCCCGCTCC 1920  
Db 6861 CCTTCCACATCAGAGAGAGCGGCTAAGGCGAGCTGGGCGAGGGGCTTCCCGCTCC 6920  
Qy 1921 TTTGGCCAGCTTTCAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980  
Db 6921 TTTGGCCAGCTTTCAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6980  
Qy 1981 CAAATGACTTCCAGAGCGCTCATCGAGGCGCAACCTCTGTGGCGGCTGAAGATG 2040  
Db 6981 CAAATGACTTCCAGAGCGCTCATCGAGGCGCAACCTCTGTGGCGGCTGAAGATG 7040  
Qy 2041 GGGCGGAGCAATTAACCGGCTGGAGTCAAGAAACAAGTGAATCTTGAATCTTTGAC 2100  
Db 7041 GGGCGGAGCAATTAACCGGCTGGAGTCAAGAAACAAGTGAATCTTGAATCTTTGAC 7100

Qy 2101 CCGCTCCAGCGGAGAGATGAGCGGAAAGTGTCCGTCCCGGCGAGATCTTGGGAAA 2160  
Db 7101 CCGCTCCCGGCGGAGAGATGAGCGGAAATATCCGTTCCGCGCGAGATCTTGGGAAA 7160  
Qy 2161 TCCAAAGAAATTTCCACAGCGATGCCCCGATGGGCACTCCCGGATTAACAACCTCCGCTG 2220  
Db 7161 CCAGAGAGTTTCCCCCAGCGTGGCCATATGGGAGCGCCCGGATTAACAACCTCCAGCTG 7220  
Qy 2221 CTGAGTCCGTGAAGGCGCGGAGCTAAGTCCCTCCAGTGGTATCATGGGGGCCCATGGCA 2280  
Db 7221 CTGAGTCCGTGAAGGAGCCGAGCTAAGTCCCTCCAGTGGTATCATGGGGGCCCATGGCA 7280  
Qy 2281 CCTACTAAGACCCCTCTTATACACTCCACGAGAGAGAGAGAGAGTGTGTGACAGA 2340  
Db 7281 CCTACTAAGAGGCGCTCATATACACTCCACGAGAGAGAGAGAGAGTGTGTGACAGAG 7340  
Qy 2341 TCCACCGTGTCTTGTGCGCTGGCGGAGCTTGGCAAAAGGCTTTGGTATGCTTCCAGACG 2400  
Db 7341 TCCACCGTGTCTTGTGCGCGGAGCTTGGCTAAGACCTTTGGCAGCTTCCGAGTGC 7400  
Qy 2401 TGGGCGGTGAGAGCGGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2460  
Db 7401 TGGGCGGCGGAGTGGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7460  
Qy 2461 GCAGATCTGAGCTGAGTGTATTCCTCATGCCCCCTTGAAGGGGAGCGGGGAG 2520  
Db 7461 AAGAGATCCGAGTGTGAGTGTATCTCATATGCCCCCTCGAGGAGAGCGGGGAG 7520  
Qy 2521 CCGATCTCAGCGAGCGGCTTGTGTGTATCCGTGATGAGAGGCGGCTGAGAGCTGCTG 2580  
Db 7521 CCGATCTCAGCGAGTGGTGTGTGTGTATCCGTGAGAGAGGAGTGTGAGAGCTGCTG 7580  
Qy 2581 TGTGTGTGATGTCTTCAATGAGCAGGCGCTGTATCAGCGCATGCGCTGGAGAGAA 2640  
Db 7581 TGTGTGTGATGTCTTCAATGAGCAGGCGGCTGTATGAGCATGCGCTGGAGAGAA 7640  
Qy 2641 AGCAAGCTCCCATCAACGCGTGTGAGCACTTTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
Db 7641 AGCAAGCTCCCATCAACGCGGCTGTGAGCACTTTGCTGCTGCTGCTGCTGCTGCTG 7700  
Qy 2701 GCTTACATATCCCGAGCGGCAACCGGCGGAGAGAGTGAAGTGAAGTGAAGTGAAG 2760  
Db 7701 GCTTACATATCCCGAGCGGCAACCGGCGGAGAGAGTGAAGTGAAGTGAAGTGAAG 7760  
Qy 2761 ATCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2820  
Db 7761 ATCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7820  
Qy 2821 AAGGCTAAGCTTCTATCAGTGAAGAGCTGCAAGCTGACGCGCCCATTCGAGCAAA 2880  
Db 7821 AAGGCTAAGCTTCTATCAGTGAAGAGCTGCAAGCTGCAAGCTGACGCGCCCATTC 7880  
Qy 2881 TCTAAATTTGGCTATGAGGCGAAGAGCTGCGGACCTTATCCAGCAAGGCTATTAACAC 2940  
Db 7881 TCTAAATTTGGCTATGAGGCGAAGAGCTGCGGACCTTATCCAGCAAGGCTATTAACAC 7940  
Qy 2941 ATCCGCTCCGTGGGAGAGCTTGTGGAAGACCTGAACCAATTTGACCAACATC 3000  
Db 7941 ATCCGCTCCGTGGGAGAGCTTGTGGAAGACCTGAACCAATTTGACCAACATC 8000  
Qy 3001 ATGGCAAAAATGAGGTTTTTGTGCTCAACAGAGAGAGAGGCGGCAAGCGAGCTGCG 3060  
Db 8001 ATGGCAAAAATGAGGTTTTTGTGCTCAACAGAGAGAGAGGCGGCAAGCGAGCTGCG 8060  
Qy 3061 CTATGCTGTTCAGAGCTTGGGGGCTCGGTGTGTGAGAGAAATGAGGCTTATGAGCTG 3120  
Db 8061 CTATGCTGTTCAGAGCTTGGGGGCTCGGTGTGTGAGAGAAATGAGGCTTATGAGCTG 8120  
Qy 3121 GTTTCACCGTCCCTGAGGCTGAGTGGGCTCGTGAAGGAGTTCAGATTCCTGCTGGA 3180  
Db 8121 GTTTCACCGTTCCTGAGGCGTGAAGGCTCTCATACGAGTTCAGATTCCTGCTGGA 8180



QY 3181 CACGGGTCGAGTTCCTGTGTAAGCGCTGGAAATCAAGAAGACCCCTATGGCTTGCA 3240  
 |||||  
 Db 8181 CACGGGTCGAGTTCCTGTGTAAGCGCTGGAAATCAAGAAGACCCCTATGGCTTGCA 8240  
 QY 3241 TATGACACCCGCTGTTTGAAGTCAAGTCACTGAGAAATACATCCGTGTAGAGAGTCA 3300  
 |||||  
 Db 8241 TATGACACCCGCTGTTTGAAGTCAAGTCACTGAGAAATACATCCGTGTAGAGAGTCA 8300  
 QY 3301 ATTATACAAATGTTGATCTGGCCCCCGAAGCCAGAGCCATTAAGTGTGCTACAGAG 3360  
 |||||  
 Db 8301 ATTATACAAATGTTGATCTGGCCCCCGAAGCCAGAGCCATTAAGTGTGCTACAGAG 8360  
 QY 3361 CGGCTTATATCGGGGGTCCCTGACTAATTTCAAAAGGAGAACTGGCTATCGCGG 3420  
 |||||  
 Db 8361 CGGCTTATATCGGGGGTCCCTGACTAATTTCAAAAGGAGAACTGGCTATCGCGG 8420  
 QY 3421 TGGCGCGGAGCGGCGTGTGAGAGTGTGCGGTAAATACCTGACATGTTAATTGAG 3480  
 |||||  
 Db 8421 TGGCGCGGAGCGGAGTGTGAGAGTGTGCGGTAAATACCTGACATGTTAATTGAG 8480  
 QY 3481 GCCTTGTGACGCTGTGAGAGTGTGCAAGAGTCCAGAGCTGACATGTCTGTGTGGAGAC 3540  
 |||||  
 Db 8481 GCCTTGTGAGCTGTGAGAGTGTGCAAGAGTCCAGAGCTGACATGTCTGTGTAAAGAGAC 8540  
 QY 3541 GGCCTTGTGCTTATCTGTGAGAGCGGGAACCCAGAGAGACGCGGAGAGCTTACAGAGTC 3600  
 |||||  
 Db 8541 GACCTTGTCTTATCTGTGAGAGCGGGAACCCAGAGAGATGCGGAGAGCTTACAGAGTC 8600  
 QY 3601 TTCACGAGAGCTATGACTAGTACTGTGCCCCCGGGGAGACCGCCCAACAGAAATAC 3660  
 |||||  
 Db 8601 TTCACGAGAGCTATGACTAGTACTGTGCCCCCGGGGAGACCGCCCAACAGAAATAC 8660  
 QY 3661 GACCTGAGAGTGTATATCATATATGTCTCTCCATGTGTGTGCGTGTGAGACATGTGATCTGCG 3720  
 |||||  
 Db 8661 GACCTGAGAGTGTATATCATATATGTCTCTCCATGTGTGTGCGTGTGAGACATGTGATCTGCG 8720  
 QY 3721 AAAAGGTATATACCTCACCCGTAACC 3749  
 |||||  
 Db 8721 AAAAGGTATATACCTCACCCGTAACC 8749  
 RESULT 11  
 AR488101  
 LOCUS AR488101 8642 bp DNA linear PAT 15-MAY-2004  
 DEFINITION Sequence 2 from patent US 6706874.  
 ACCESSION AR488101  
 VERSION AR488101.1 GI:47253849  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 8642)  
 Kukulj,G. and Pause,A.  
 TITLE Self-replicating RNA molecule from hepatitis C virus  
 JOURNAL Patent: US 6706874-A 2 16-MAR-2004;  
 FEATURES  
 location/Qualifiers  
 1..8642  
 /organism="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN  
 Query Match 87.8%; Score 3292.2; DB 6; Length 8642;  
 Best Local Similarity 92.3%; Pred. No. 0;  
 Matches 3462; Conservative 2; Mismatches 285; Indels 0; Gaps 0;  
 QY 1 TGGGAGAGGAGCTTTCACAGGCTCACCCAGTGGATGCCACTTCTGTCCAAACAAG 60  
 |||||  
 Db 4046 TGGGAGAGGAGCTTTCACAGGCTCACCCAGTGGATGCCACTTCTGTCCAAACAAG 4105  
 QY 61 CAGGACGAGAGCAACTTCCCTTACTGTGTGCGGTACAGGAGTACTGTGTGCGGTAGG 120  
 |||||  
 Db 4106 CAGGACGAGAGCAACTTCCCTTACTGTGTGCGGTACAGGAGTACTGTGTGCGGTAGG 1260  
 |||||

QY 121 CAGGCCCACTTCATCATGGATCAAAATGTGAAGTGTCTATACGGCTAAAGCCTACT 180  
 |||||  
 Db 4166 CAGGCTCCACTCATGTGTGGACCAAAATGTGAAGTGTCTATACGGCTAAAGCCTACT 4225  
 QY 181 CTGCGGGGCGCAACACCTTGTGTATAGGCTGGAGCGCTCCAAAAGAGTCAACCTTC 240  
 |||||  
 Db 4226 CTGCAAGGGCGCAACGCGCCCTGTGTATAGGCTGGAGCGCTCCAAAAGAGTCAACCTTC 4285  
 QY 241 ACACACCCCAATACCAAAATTCATATGAGCATGTGATGACGCGACCTGGAGGTGTACG 300  
 |||||  
 Db 4286 ACACACCCCAATACCAAAATTCATATGAGCATGTGATGACGCGACCTGGAGGTGTACG 4345  
 QY 301 AGCAGCTGGGTGTGTGTGGCGGGGTCCTTGACAGCTCTGGCTGCTATTTGCTTGACACA 360  
 |||||  
 Db 4346 AGCAGCTGGGTGTGTGTGGCGGGGTCCTTGACAGCTCTGGCTGCTATTTGCTTGACACA 4405  
 QY 361 GGCAGCGTGTCTATTTGGGTAGAGATCATTTGTCCGGGCGGCGGCTATTTGTTCCCGAC 420  
 |||||  
 Db 4406 GGCAGCGTGTCTATTTGGGTAGAGATCATTTGTCCGGGCGGCGGCTATTTGTTCCCGAC 4465  
 QY 421 AGGGAAGTCTCTTACCAAGAGTTCATGAGATGGAAGAGTGGCGGCTGCACTCCCTTAC 480  
 |||||  
 Db 4466 AGGGAAGTCTCTTACCAAGAGTTCATGAGATGGAAGAGTGGCGGCTGCACTCCCTTAC 4525  
 QY 481 ATCAGAGAGGGAATGACAGTGTGCGGAGAGTTCAGAGAAACGCTCGGAGTGTCTGAC 540  
 |||||  
 Db 4526 ATCAGAGAGGGAATGACAGTGTGCGGAGAGTTCAGAGAAACGCTCGGAGTGTCTGAC 4585  
 QY 541 ACAGCCACCAAGCAAGCGAGGCGGCTGCTCCGCTGGTGGAGTCCAAGTGGAGACCCCTT 600  
 |||||  
 Db 4586 ACAGCCACCAAGCAAGCGAGGCGGCTGCTCCGCTGGTGGAGTCCAAGTGGAGACCCCTT 4645  
 QY 601 GAAACCTTCTGGGCGAAACATGTGGAATTATACGCGGATTAACATCTTAAGACAGG 660  
 |||||  
 Db 4646 GAAACCTTCTGGGCGAAACATGTGGAATTATACGCGGATTAACATCTTAAGACAGG 4705  
 QY 661 TTGTCACTGTGCTGGGAAATCCCGGATTCATCATGTGAGCTTACAGGCTCTGTG 720  
 |||||  
 Db 4706 TTGTCACTGTGCTGGGAAATCCCGGATTCATCATGTGAGCTTACAGGCTCTGTG 4765  
 QY 721 ACTAGCCGCTCAACCAACCAATCTACCTCTGCTTTAACTCTGCGGGAGATGGGTAGCC 780  
 |||||  
 Db 4766 ACTAGCCGCTCAACCAACCAATCTACCTCTGCTTTAACTCTGCGGGAGATGGGTAGCC 4825  
 QY 781 GCCCAACTGTCTCCCCCAAGTGTCTTCAAGCTTTGTGTAGGCGCGGCAATGCTGTGTCG 840  
 |||||  
 Db 4826 GCCCAACTGTCTCCCCCAAGTGTCTTCAAGCTTTGTGTAGGCGCGGCAATGCTGTGTCG 4885  
 QY 841 GCTGTGGAGCATATAGGCTTGGGAAGTGTCTGTGGAACATCTTGGCGGGCTATGAGCA 900  
 |||||  
 Db 4886 GCTGTGGAGCATATAGGCTTGGGAAGTGTCTGTGGAACATCTTGGCGGGCTATGAGCA 4945  
 QY 901 GGAAGTGGACAGCGCGCTGTGAGCTTTAAGTGTATGAGCGGAGAAATGCCCTTCACCGAG 960  
 |||||  
 Db 4946 GGAAGTGGACAGCGCGCTGTGAGCTTTAAGTGTATGAGCGGAGAAATGCCCTTCACCGAG 5005  
 QY 961 GACCTGTATTAATTAATCTTCCCTGCAATCTCTCTCTGTGTGTGTCCTGCTGTGCGGGTGTG 1020  
 |||||  
 Db 5006 GACCTGTATTAATTAATCTTCCCTGCAATCTCTCTCTGTGTGTGTCCTGCTGTGCGGGTGTG 5065  
 QY 1021 TGGCAGAGCATATGAGTGTGGAGCGAGTGGGTCAGAGGAGAGGAGGCTGTGAGTGAATGAC 1080  
 |||||  
 Db 5066 TGGCAGAGCATATGAGTGTGGAGCGAGTGGGTCAGAGGAGAGGAGGCTGTGAGTGAATGAC 5125  
 QY 1081 CGGCTATATAGCTTGTGCTGTGCGGGGTAAACATGTTTCCCAACGACATATGTGCCAGAG 1140  
 |||||  
 Db 5126 CGGCTATATAGCTTGTGCTGTGCGGGGTAAACATGTTTCCCAACGACATATGTGCCAGAG 5185  
 QY 1141 AGGAGAGCGGAGACAGTGTCACTAGATTCCTTCCGACCTTATATACCCAACTGTG 1200  
 |||||  
 Db 5186 AGGAGAGCGGAGACAGTGTCACTAGATTCCTTCCGACCTTATATACCCAACTGTG 5245  
 QY 1201 AAGAGGCTCCACAGTGAATTAACAGAGACTGTCTCACGCGCTGTCTCGGCTGTGAGCTTA 1260  
 |||||

|||||  
Db 5246 AAGAGCTTCACGAGTGAATCAACGAGACTGCTCCAGCTCCGCTGCTGCTGA 5305  
Qy 1261 AGGAGATGTTGGGACGAGATATGACAGATTTGGCTGACTTGAAGACCTGGCTCAAGTCC 1320  
Db 5306 AGGAGATGTTGGGATGATGACAGATTTGATGATTTCAAGACCTGGCTCAAGTCC 5365  
Qy 1321 AAGCTCTGCGGAGATTTACCGGAGATCCCTTTTTCATGACCAAGTGGTGAACAAGGG 1380  
Db 5366 AAGCTCTGCGGAGATTTGCGGGAGTCCCTTCTTCATGTCACAAGTGGTGAACAAGGG 5425  
Qy 1381 GTTGGCGGGGAGACGGCATATGACAGACCACTGCTCATGTTGAGAGACAGATCACCGGA 1440  
Db 5426 GTCTGGCGGGGACCGCATATGACAAACCTGCCATGTGAGAGACAGATCACCGGA 5485  
Qy 1441 CATGTCAAAAAGGTTCCATGAGGATGTGGGCTTAAGACCTGTAGTAACATGTGGCAT 1500  
Db 5486 CATGTCAAAAAGTTCATGAGGATGTGGGCTTAAGACCTGTAGTAACATGTGGCAT 5545  
Qy 1501 GGAACATTTCCCATCAACGATACACCAAGGAGCCCTGACAGCCCTCCCGAGGCCAAG 1560  
Db 5546 GGAACATTTCCCATCAACGATACACCAAGGAGCCCTGACAGCCCTCCCGAGGCCAAG 5605  
Qy 1561 TATTCAAGGAGCGCTGTGGCGGATGTGCTGTGAGAGATGCTGAGATTAACGCGGATGGG 1620  
Db 5606 TATTCAAGGAGCGCTGTGGCGGATGTGCTGTGAGAGATGCTGAGATTAACGCGGATGGG 5665  
Qy 1621 GATTTCCACTAGTGAAGACATGACCACTGACCAAGTAAATGCGGCTCCAGGTTCCA 1680  
Db 5666 GATTTCCACTAGTGAAGGACATGACCACTGACCAAGTAAATGCGGCTCCAGGTTCCG 5725  
Qy 1681 GCCCCGGAATTTCTTCACAGAGTGAATGGGATGGGCTGACACAGTACGCTCCGCGCTGC 1740  
Db 5726 GCCCCGGAATTTCTTCACAGAGTGAATGGGATGGGCTGACACAGTACGCTCCAGGCTGC 5785  
Qy 1741 AAACTCTCTTACGAGGAGAGTTCATTCAGGTCGGGCTCAACCAATACCTGTGGG 1800  
Db 5786 AAACTCTCTTACGAGGAGAGTTCATTCAGGTCGGGCTCAACCAATACCTGTGTGGG 5845  
Qy 1801 TCCGAGTCCGATGCGAGCCGGAACCGGATGTAGCAATGCTCACTTCATGACTCACGAC 1860  
Db 5846 TCCGAGTCCGATGCGAGCCGGAACCGGATGTAGCAATGCTCACTTCATGACTCACGAC 5905  
Qy 1861 CCTTCCACATCAACGAGAGACGAGCTAAGCGAGCTGGCCAGGAGGATCTCCCTCC 1920  
Db 5906 CCTTCCACATTAACGAGAGACGAGCTAAGCGATAGCTGGCCAGGAGGATCTCCCTCC 5965  
Qy 1921 TTGGCAGCTCTTCACTAGCAAGTGTCTGCGCTTCTGACAGGCGCAATACATTAAC 1980  
Db 5966 TTGGCAGCTCTTCACTAGCAAGTGTCTGCGCTTCTGAAAGGCGCAATACATTAAC 6025  
Qy 1981 CAAATGATCTTCCAGAGCGTCACTGATGAGGCGCAACCTCCGATGGCGGCAATGAGTG 2040  
Db 6026 CGTATATATCTCCCGAGCGTGAACCTATGAGGCGCAACCTCCGATGGCGGCAATGAGTG 6085  
Qy 2041 GCGCGGAGCATTAACCGCTGAGAGTCAAGAAACAAGTAAATCTGGAATCTTTTGAC 2100  
Db 6086 GCGCGGAGCATTAACCGCTGAGAGTCAAGAAACAAGTAAATTTTGGACTCTTTTGAG 6145  
Qy 2101 CCGCTCGAGCGAGAGATGAGCGGAAAGTGTCCGTCCCGCGAGATCTTCGGAAA 2160  
Db 6146 CCGCTCGAGCGAGAGATGAGCGGAAAGTGTCCGTCCCGCGAGATCTTCGGAGG 6205  
Qy 2161 TCCAGAAATTTCCACAGAGAGATGCGGAGAGGCGAGGCTTCAACCCCTCGGCTG 2220  
Db 6206 TCCAGAAATTTCTTGAAGAGATGCCATATGAGGCGAGCGCGGATTAACAACCTTCAC 6265  
Qy 2221 CTGAGATCTTGAAGAGCGCGAGCTAGCTCCCTCAAGTGTACATGAGTGGCGCACTGCA 2280  
Db 6266 TTGAGATCTTGAAGAGCGCGAGCTAGCTCCCTCAAGTGTACATGAGTGGCGCACTGCG 6325  
Qy 2281 CTTACTAAGACCTCTCTATACACTTTCACGAGAGAGAGACGTTTCTGACAGAA 2340  
|||||

Db 6326 CTGCGAAGCCCTTCGATACCACTCCAGGAGAGAGACGTTGTCTGTAGAA 6385  
Qy 2341 TCCACCGATCTTCTGCGCTTGGCGAGAGTGTGCGACAAAGCTTTCGATGCTCCGAACG 2400  
Db 6386 TCTACCGATCTTCTGCGCTTGGCGAGAGTGTGCGACAAAGACCTTGGGAGCTCCGAATCG 6445  
Qy 2401 TGGCGCTTGAACGCGGACCGGCAACCGCCCTCTGACCAACCTTCCGACGAGCGGGA 2460  
Db 6446 TGGCGCTTGAACGCGGACCGGCAACCGCCCTCTGACCAACCGCCCTCCAGACGAGCGAG 6505  
Qy 2461 GAGAGATCTGAGGATGAGTGTATCTGCATGAGCCCTTGGAGGAGGAGCGGAGGAG 2520  
Db 6506 GAGAGATCTGAGGATGAGTGTATCTGCATGAGCCCTTGGAGGAGGAGCGGAGGAG 6565  
Qy 2521 CCGATCTGAGGAGCGGATCTTGTGTACCTGTAGTGAAGAGCGGATGAGAGAGTGTGTG 2580  
Db 6566 CCGATCTGAGGAGCGGATCTTGTGTACCTGTAGTGAAGAGCGGATGAGAGAGTGTGTG 6625  
Qy 2581 TGTGCTGATGTCTTACATGAGACAGGCGCTGTATCAACGCTATGCGTGGAGGAA 2640  
Db 6626 TGTGCTGATGTCTTACATGAGACAGGCGCTGTATCAACGCTATGCGTGGAGGAA 6685  
Qy 2641 AGGAGCTGCCATCAACGCTTGAAGCACTCTTGTGCGGTGACCAACATGTGTCTAC 2700  
Db 6686 ACCAGCTGCCATCAATGATGACTGAGCAACTCTTGTGCGGTGACCAACATGTGTCTAC 6745  
Qy 2701 GCTACCAATCCCGACGCGCAAGCCGAGCGGCAAGAGGTCACCTTTGACAGACTGCA 2760  
Db 6746 GCTACCAATCTTCGAGCGCGCAAGCCGAGCGGCAAGAGGTCACCTTTGACAGACTGCA 6805  
Qy 2761 ATCTGAGAGATCACTACAGAGAGTGTCTGAGAGATGAGGCGGAGGCTTCACAGTT 2820  
Db 6806 ATCTGAGAGAGATCACTACAGAGAGTGTCTGAGAGATGAGGCGGAGGCTTCACAGTT 6865  
Qy 2821 AAGGCTAACTTCTATCAAGAGAGAGCTGCAAGCTGACGCGCCACATTTGGGCAAA 2880  
Db 6866 AAGGCTAACTTCTATCAAGAGAGAGCTGCAAGAGCTGCAAGCTGACGCGCCACATTTGGGCAAG 6925  
Qy 2881 TCTAAATTTGGCTATGAGGAGAGAGCTGCGGAACTATCCAGAGGCGCATTAACAC 2940  
Db 6926 TCTAAATTTGGCTATGAGGAGAGAGCTGCGGAACTATCCAGAGGCGCATTAACAC 6985  
Qy 2941 ATCCGCTCGGTGGAGAGACTTGTGGAAGACCTGGAACACCAATTGACACCACTATC 3000  
Db 6986 ATCCGCTCGGTGGAGAGACTTGTGGAAGACCTGGAAGACCAATTGACACCACTATC 7045  
Qy 3001 ATGGCAAAAATGAGGTTTTCTGCGTCAACCAAGAGAGAGGAGCGGCAAGCGCTGCG 3060  
Db 7046 ATGGCAAAAATGAGGTTTTCTGCGTCAACCAAGAGAGAGGAGCGGCAAGCGCTGCG 7105  
Qy 3061 CTTATCGTGTCCAGACTTGGGAGTCCGTGTGTGAGAGAAATGAGCCCTTATGACGTG 3120  
Db 7106 CTTATCGTGTCCAGACTTGGGAGTCCGTGTGTGAGAGAAATGAGCCCTTATGACGTG 7165  
Qy 3121 GTCTCAACCTTCCCTCAAGCTGTGATGAGGCTCTCTGATGAGATTCAGTATTTCTGGA 3180  
Db 7166 GTCTCAACCTTCCCTCAAGGCGGTGATGAGGCTCTCTGATGAGATTCAGTATTTCTGGA 7225  
Qy 3181 CAGCGGATGAGTCTCTGTGGAACGCTGGAATTCAAAGAGACCCCTATGGGCTTTGCA 3240  
Db 7226 CAGCGGATGAGTCTCTGTGGAATGCTGGAAGGCGAAGAAATGCTTATGGGCTTTGCA 7285  
Qy 3241 TATGACACCGGCTGTTTGACTCAAGGATCACTGAGATGACATCCGTGTGAGAGTCA 3300  
Db 7286 TATGACACCGGCTGTTTGACTCAAGGATCACTGAGATGACATCCGTGTGAGAGTCA 7345  
Qy 3301 ATTATCAATGTTTGAATTTGAGCTTGGCGCCCGAAGCGAGACAGGCTTAAGTGTCTCAAG 3360  
Db 7346 ATCTACCAATGTTTGAATTTGAGCTTGGCGCCCGAAGCGAGACAGGCTTAAGTGTCTCAAG 7405  
Qy 3361 CGGCTTTATATCGGAGGATCCCTGATCAATTAATCAAAAAGGAGAGACTGCGGCTATCGCGG 3420  
Db 7406 CGGCTTTATATCGGAGGATCCCTGATCAATTAATCAAAAAGGAGAGACTGCGGCTATCGCGG 7465  
|||||



QY 3421 TGGCGCGGAGCGCGCTGTGACGACTAGCTGGCGTATACCTCACATGTTACTTGAAG 3480  
 DB 7466 TGGCGCGGAGCGCGCTGTGACGACGACGCTGGCGTATACCTCACATGTTACTTGAAG 7525  
 QY 3481 GCGCTGTGACGCGCTGTGACGAGCTGCAAAAGCTCCAGACTGACGATGCTGTGTGCGGAGAC 3540  
 DB 7526 GCGCGCTGTGACGCGCTGTGACGAGCTGCAAAAGCTCCAGACTGACGATGCTGTGTGCGGAGAC 7585  
 QY 3541 GCGCTGTGCTTATCTGTGAGACGCGGAGAACCGAGAGACGCGGCGGACCTTACCGAGTC 3600  
 DB 7586 GACCTGTGCTTATCTGTGAGAGCGCGGAGAACCGAGAGACGCGGCGGACCTTACCGAGTC 7645  
 QY 3601 TTCACGAGAGCTATGACGACTAGTACTGTGCCCCCGCGGAGACCGCGCCCAACGAGATAC 3660  
 DB 7646 TTCACGAGAGCTATGACGACTAGTACTGTGCCCCCGCGGAGACCGCGCCCAACGAGATAC 7705  
 QY 3661 GACCTGAGATTGATTAACATCATCTCTCTCAATGTGTGCGTGGCGACGATGATCTGGC 3720  
 DB 7706 GACCTGAGATTGATTAACATCATCTCTCTCAATGTGTGCGTGGCGACGATGATCTGGC 7765  
 QY 3721 AAAAGGATATCTACCTCACCGCTGACCC 3749  
 DB 7766 AAAAGGATATCTACCTCACCGCTGACCC 7794  
 RESULT 12  
 AX472292  
 LOCUS AX472292 8642 bp DNA linear PAT 09-AUG-2002  
 DEFINITION Sequence 2 from Patent WO2052015.  
 ACCESSION AX472292  
 VERSION AX472292.1 GI:22207307  
 KEYWORDS  
 SOURCE Hepatitis C virus  
 ORGANISM Hepatitis C virus  
 1. Hepatitis C virus  
 Hepatitis C virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.  
 1. Kukoji, G. and Pause, A.  
 Self-replicating RNA molecule from hepatitis C virus  
 Patent: WO 02052015-A 2 04-JUL-2002;  
 BOEHRINGER INGELHEIM (CANADA) LTD. (CA)  
 Location/Qualifiers  
 1..8642  
 /organism="Hepatitis C virus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:11103"  
 1802..8407  
 /note="unassigned protein product"  
 /codon\_start=1  
 /protein\_id="CAD43568.1"  
 /db\_xref="GI:22207308"  
 /translation="MDREMAASCGAVFVGLILLLISPHYKFLAELIWLQYFTRTA  
 EAHLYQVPLNVRGRDAVILITCAIHELIPTIKILALIGPLMVLQAGITKPY  
 FPAHGLIRACMLVRKAVGGHYOMALALITGVYVDHLITPDMAGKATLAV  
 AVEPVVSDMETKVTIVTACGADTAACDIIILGIVSARGRHILGPDMLGCGWRLA  
 PITAYSOQTRGLGCTITSLTGRDNQVESEVYSTATDSFLATVNGVCTVTHGA  
 GSTITLGPGLITQMTNVDQILVGMQAPGARSLPCTGSSDLVLRHADVIR  
 RGDNRGLSPRPVSLKSGSGPLCEGDAVGI FRAVCTRGVAKVDVFPVSM  
 ETTMRSPVFTDNRSPAVPOTFOVAHLHAPTSKSTKPAAGAGYKVLVNPSSA  
 ATLFGAAMSKAHGDIPNIRTVRTITGATPSTYTGKLAAGSGGAYDIIICDE  
 CHSDSTTLTGITVDLQAEATAGARLVLATATPAGSVTPPHNIEVLSNGELPF  
 YKATIPETIKGRHLIFCHSKKCBELAKLSGLGNNAVAYIRGIDVSIPTSGVIT  
 VVATDALMTGFTGDPDSVIDCNTVQTVDPSLDPTFTLETTVPDVAIRSQRRT  
 RGRMGIYRFVPEGRSGMFDPSVLCEDGACAVELIPATSVRLAYLNTPLP  
 VCODHLFEWESVFTGLTHIDAPLSQKQDNFPLYVAQATVCAARAQAPSPMDM  
 WCLILKPLTLPGLTLYRLAGVONEVITTHBITKYIACMSADLEVSTVTLVGG  
 VLAALAYCLTGSVIVIGRIILSGKPAIIPREVLVYREDEMEGASHLPYEOGMO  
 LAQFKQKAIIGLQITAKQAAAPVSESKWRTLEFWAKMNPISGIOYLAGSTL  
 PGNPAIASHMAFTASTITPLTQHTLLPILNGMVAQAQAPSAAPILBPGALVAV  
 GSGLGKVLVDITLAGAGVAGALVAFKWSGEMSPTEBDVNLPLILBPGALVAV  
 CAALIRRHVPGGAVOMNRILAFASRGHVSPTHYVESDAAAVTOLISLITTO  
 LKRLHQMINEDESTPCSGMWLADVMDICTVLTIDPKTWLQSKLLPRLGVPPFSQR

GYKVRGDIQIMOTPCGAOITGVHNCMSMRIVGRBTCNSMTWHTFPINAYTTGPTCT  
 PSPAPVSRALRVAAEYVEYTRVVDPFYVTMTDNYKPCQVAPAEFTVDGR  
 LHRVAPACKPLREBEYTPVLGNQYLVGSOLEPEEPDVAVLITSLMIDPSHITAEAK  
 RLRAASPSRSLSSASSQASABIKTCTTRHDSPADLILENLIRQEGNITVE  
 SERKVIILDSFEPLOAEDEREVSVPAELIIRSRKPPRAMPITARDYINPGLIESKD  
 PDVPPVHVGCPLPPAKAPP1PEPRKRTVIVSESVSSALAEALAKTGVSSSSAVD  
 SGTAASPPQPSDDGAGSDVESISMPLEBPGDPDLSGDSWSTVSEBSIDVVC  
 SMSTWTGALITPCAEETKLPINALSNBLRHHNLVYATTSASLSLROKRTFDRLO  
 YLDDHYRDLKEMKASTVYKAKLSVEEAKCLTPHARSARXFGVAGKADRNLSKVA  
 NHRISWKULDETEPIDTTIMAKREVCVQPEKGRPARLIVPDLGVRCEKKA  
 LIDVSTLPOAVWGSYGFQYSPGQVEFLVNAWKAKCPMGFAIDTRCEPDLSTEND  
 IIRVESIYQCCDLAPBARQALIRSLTERLYIGPLTYSKQNCGRARCRASGLVITSCG  
 NTLTCYLAKAARAKLQDCTMLVCGDDLVJGCSAGQOEDASIRAFETARVSA  
 PRDPEKPYDLELITSCSNYSVAHDGKRYVYLRDPTPLAAARETARVFN  
 SMLGNTIMAPITLMAEMITHEFTSLILAOBLERKLDQOIGACVSIIEPLDIPOLIO  
 RHGLSAPFLSHSISPEINRVASCLBKGLVPPRLVRHARSVRALLISQGGRAATCG  
 KILFMAVRKCLKLPIPAASQULSSWFAVAGSDIYHSIRARPRFMWCLLILS  
 VGVILLPNR"  
 4446  
 /note="r = a or g"  
 6268  
 /note="r = a or g"  
 variation  
 variation  
 ORIGIN  
 Query Match 87.8%; Score 3292.2; DB 6; Length 8642;  
 Best Local Similarity 92.3%; Pred. No. 0;  
 Matches 3462; Conservative 2; Mismatches 285; Indels 0; Gaps 0;  
 QY 1 TGGAGAGGCGCTTTCACAGGCTCACCCAGTGATGCCACTTCTGTCCAAACAAG 60  
 DB 4046 TGGAGAGGCGCTTTCACAGGCTCACCCAGTGATGCCACTTCTGTCCAAACAAG 4105  
 QY 61 CAGGACGAGACAACTTCCCTTACCTGTGAGCGTACCAAGCTTCTGTGCGTAAAGCC 120  
 DB 4106 CAGGACGAGACAACTTCCCTTACCTGTGAGCGTACCAAGCTTCTGTGCGTAAAGCC 4165  
 QY 121 CAGGCCCACTCATCATGTGAGTCAATGTGGAAGTGTCTATAAGCTTAAAGCC 180  
 DB 4166 CAGGCTTCACCTTCATGTGAGACCAATGTGGAAGTGTCTATAAGCTTAAAGCC 4225  
 QY 181 CTGCGGCGGCGCAACACCTTGTCTGTATAGCTGTGAGAGCGGTCCAAAGAGAGTCAAC 240  
 DB 4226 CTGCAAGGCGCAACCGCCCTGTGTATAGCTGTGAGAGCGGTCCAAAGAGAGTCAAC 4285  
 QY 241 ACACACCCCATTAACCAATTCATCATGTGAGTCAAGCGGACCTGAGAGTCTGACAG 300  
 DB 4286 ACACACCCCATTAACCAATTCATCATGTGAGTCAAGCGGACCTGAGAGTCTGACAG 4345  
 QY 301 AGCAGCTGGGTGTGTGAGCGGCGGCTTTCAGCTCTGAGCTGCTTATGCTTGAACA 360  
 DB 4346 AGCAGCTGGGTGTGTGAGCGGCGGCTTTCAGCTCTGAGCTGCTTATGCTTGAACA 4405  
 QY 361 GGCAGCGTGTGATTTGTGAGTATGAGTCACTTGTCCGAGCGGCGCTATTGTTCCGAG 420  
 DB 4406 GGCAGCGTGTGATTTGTGAGTATGAGTCACTTGTCCGAGCGGCGCTATTGTTCCGAG 4465  
 QY 421 AGGGAAGTCTCTTACCAAGAGTTCATGATGAGGAAGTGGCGTCTGCACTTCCCTTAC 480  
 DB 4466 AGGGAAGTCTCTTACCAAGAGTTCATGATGAGGAAGTGGCGTCTGCACTTCCCTTAC 4525  
 QY 481 ATGAGACAGGGAATGAGCTCGCGGAGCTTCAAGCAAAAAGCGCTCGGGTTGCTGAG 540  
 DB 4526 ATGAGACAGGGAATGAGCTCGCGGAGCTTCAAGCAAAAAGCGCTCGGGTTGCTGAG 4585  
 QY 541 ACAGCCACCAAGCAAGCGAGGCGCTGCTCCCGTGTGAGTCCAAAGTGGCGAGCCCTT 600  
 DB 4586 ACAGCCACCAAGCAAGCGAGGCGCTGCTCCCGTGTGAGTCCAAAGTGGCGAGCCCTT 4645  
 QY 601 GAGACTTCTGGCGCAACATGTGGAACCTTATCAGCGGAGTACAGTACTTGAAGAGC 660  
 DB 4646 GAGACTTCTGGCGCAAGATATGTGGAATTTTATCAGCGGAGTACAGTACTTGAAGAGC 4705  
 QY 661 TTGTCACTCTGCTGGAATCCCGGAGTGTGATCATGATGAGCGGTTCACAGCTCTGTC 720

|||||  
Db 4706 TTGTCACTCTGCTGCGCAACCCCGCATGCACTGATGCAATTCACAGCTCTATC 4765  
QY 721 ACTAGCCCGCTCAACCAATCACTCTGCTTAAATCCTGGGGGGATGGGTAGCC 780  
Db 4766 ACCAGCCCGCTCAACCAATCACTCTGCTTAAATCCTGGGGGGATGGGTAGCC 4825  
QY 781 GCCCACTGCTCCCGCACTGCTTCACTTTCGTAGCGCCGCAATTCGTGTGCG 840  
Db 4826 GCCCACTGCTCTCCCAAGCGCTGTGCTTGTGTAGCGCCGCAATCGCTGGAGCG 4885  
QY 841 GCTGTGGCAGCATAGGCTTTGGGAAGTCTGTGTGACATCTTGGCGGGCTATGAGCA 900  
Db 4886 GCTGTGGCAGCATAGGCTTTGGGAAGTCTGTGTGATATTTTGGCAGGTATGAGCA 4945  
QY 901 GGAGTGGCAGGCGCGCTGCGCTTTAAGTCAATGAGCGGCAAAATGCCCTCACCGAG 960  
Db 4946 GGGGTGGCAGGCGCGCTGCGCTTTAAGTCAATGAGCGGCAAAATGCCCTCACCGAG 5005  
QY 961 GACCTGGTTAACTTAATCTCCCTGCAATCCTCTCTGTCGTCCTGTCGTGCGGAGTCGTG 1020  
Db 5006 GACCTGGTTAACTTAATCTCCCTGCAATCCTCTCTGTCGTCCTGTCGTGCGGAGTCGTG 5065  
QY 1021 TGGCAGCGCATCTGCTGCGCAAGTGGGTCAGGGGAGGGGCTGTGCACTGATGAAAC 1080  
Db 5066 TGGCAGCGCATCTGCTGCGCAAGTGGGTCAGGGGAGGGGCTGTGCACTGATGAAAC 5125  
QY 1081 CGGCTATATGCGTTGCGCTGCGGGGGTAAACATGTTTCCGCCAGCACTATGTCAGAG 1140  
Db 5126 CGGCTATATGCGTTGCGCTGCGGGGGTAAACATGTTTCCGCCAGCACTATGTCAGAG 5185  
QY 1141 AGCGAGCGCGCAGCAGTGTCACTAGATCCTCTCCGACCTTAATCAACCACTGTG 1200  
Db 5186 AGCGAGCGCGCAGCAGTGTCACTAGATCCTCTCTGTCATGTCCTTAACATCACTCAGCTGCTG 5245  
QY 1201 AAGAGGCTCCACCACTGATGATTAACAGAGACTGCTCCAGCCCTGCTCGGCTGTGCTGA 1260  
Db 5246 AAGAGGCTCCACCACTGATGATTAACAGAGACTGCTCCAGCCCTGCTCGGCTGTGCTGA 5305  
QY 1261 AGGAGTGTGGGACCTGGATATGACAGTGTGGCTGACTTAACAAGCTGAGCTCCAGTCC 1320  
Db 5306 AGGAGTGTGGGACCTGGATATGACAGTGTGTGACTGATTTCAAGACTGAGCTCCAGTCC 5365  
QY 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCAACCTGAGGTAACAAGGG 1380  
Db 5366 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGTAACCTGAGGTAACAAGGA 5425  
QY 1381 GTCTGCGGGGAGACGCGCATCATGACCACTGCTCATGTGAGACAGATCAACCGGA 1440  
Db 5426 GTCTGCGGGGAGACGCGCATCATGACCAACCTGCTCATGTGAGACAGATCAACCGGA 5485  
QY 1441 CATGTCAAAAACGGTTCCATGAGGATGTGTTGGGCTTAAGACCTGTATTAATCATGTGGCAT 1500  
Db 5486 CATGTCAAAAACGGTTCCATGAGGATGTGTTGGGCTTAAGACCTGTATTAATCATGTGGCAT 5545  
QY 1501 GGAACATTTCCCATCAACGATACACCAACCGGACCTCTGCAAGCCCTCCCGAGCGCAAC 1560  
Db 5546 GGAACATTTCCCATCAACGATACACCAACCGGACCTCTGCAAGCCCTCCCGAGCGCAAT 5605  
QY 1561 TATTCAGGGGCTGTGCGGGGTGTGCTGTGAGAGTACGTGAGGTTACCGGGGTGGG 1620  
Db 5606 TATTCAGGGGCTGTGCGGGGTGTGCTGTGAGAGTACGTGAGGTTACCGGGGTGGG 5665  
QY 1621 GATTTCACTACGTAAGAGATGACCACTGACGAACCTTAATAATCCCGTCCAGGTTCCA 1680  
Db 5666 GATTTCACTACGTAAGAGATGACCACTGACGAACCTTAATAATCCCGTCCAGGTTCCG 5725  
QY 1681 GCCCCGAATTTCTTCAAGAAATGATGGGATGGGCTGCAAGATGAGCTCGGCGCTGC 1740  
Db 5726 GCCCCGAATTTCTTCAAGAAATGATGGGATGGGCTGCAAGATGAGCTCGGCGCTGC 5785  
QY 1741 AAACCTCTCTACCGGAGAGATCAATTCAGAGTCCGGCTCAACCAATCTGTGTGGG 1800  
|||||

Db 5786 AAAACCTCTCTACCGGAGAGATCAATTCAGAGTCCGGCTCAACCAATCCTGTGGG 5845  
QY 1801 TGCAGACTCCATGCGAGGCCCAACCGGATGTAGAGTGTCACTTCCATGCTCACCGAC 1860  
Db 5846 TCACAGCTCCCATGCGAGGCCCAACCGGATGTAGAGTGTCACTTCCATGCTCACCGAC 5905  
QY 1861 CCTTCCCATATCAACAGAGAGACGGCTTAAGCCGAGCTTGGCCAGGGAGTCTCCCTCTCC 1920  
Db 5906 CCTTCCCATATCAACAGAGAGACGGCTTAAGCCGAGCTTGGCCAGGGAGTCTCCCTCTCC 5965  
QY 1921 TTGGCAGCTTTCACTAGCTAGCCAGTGTGTGGGCTTCTCTGAAAGGCGCATATCAATTAAC 1980  
Db 5966 TTGGCAGCTTTCACTAGCTAGCCAGCTGTGTGGGCTTCTCTGAAAGGCGCATATCAATTAAC 6025  
QY 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGGCGGACATGAGATG 2040  
Db 6026 CGTGATGACTTCCCGGAGCGCTGACCTCATGAGGCCAACCTCTGTGGCGGACATGAGATG 6085  
QY 2041 GGGGGGACATTAACCCGCGTGAATGAGAAACAAGTGTATCTTGAATCTTTTCAG 2100  
Db 6086 GGGGGGACATTAACCCGCGTGAATGAGAAACAAGTGTATTTTGAATCTTTTCAG 6145  
QY 2101 CGGCTCCGAGCGGAGGAGATGAGCGGGAAGTCCGTCGCCGCGGAGATCCTGCGGAA 2160  
Db 6146 CGGCTCAAGCGGAGGAGATGAGGGAAGTATCCGTTTCGCGGAGATCTTGCGGAGG 6205  
QY 2161 TCCAGAAATTTCCACACAGCGATGCGCGCATGGGCAAGCCGAGTTTAAACCTTCGCTG 2220  
Db 6206 TCCAGAAATTTCCCTGAGCGATGCGCATATGGGCAAGCCGAGTTTAAACCTTCAGCTG 6265  
QY 2221 CTGAGTCTTGAAGGCCCCCGGACTATGCTTCCTTCAGTGTATATGGGTGCCACTGCCA 2280  
Db 6266 TTRGAGTCTTGAAGGACC CGGACTATGCTTCCTTCAGTGTATACAGGAGTCCATTGGCCG 6325  
QY 2281 CCTACTAAGACCCCTCTTAATCACTTCAACGAGGAGAGAGACAGTGTGTTGACAGA 2340  
Db 6326 CTGCAAGGCCCTCTCGATACCACTTCAACGAGGAGAGAGAGAGTGTGCTGTAGAA 6385  
QY 2341 TCCACCGTGTCTTTCGCTTCGCGGAGTTCGCCAACAAGCTTTCGATAGCTCCGAACCG 2400  
Db 6386 TCTACCGTGTCTTTCGCTTTCGCGGAGTTCGCCAACAAGCTTTCGCGAGCTCCGATGCG 6445  
QY 2401 TCGGCGGTGACAGCGGCAACGGCAACCGCCCTCTGACCAACCTTCCGACGCGCGGA 2460  
Db 6446 TCGGCGGTGACAGCGGCAACGGCAACCGGCTCTCTGACAGGCTTCCGACGCGCGGAC 6505  
QY 2461 GCAGATCTGAAGTGAATGCTATTCCTTCATAGCCCCCTTGAAGGGGGAGCGGGGGAG 2520  
Db 6506 GCAGATCTGAAGTGAATGCTATTCCTTCATAGCCCCCTTGAAGGGGGAGCGGGGGAG 6565  
QY 2521 CCGGATCTGACGACCGGCTTTCGCTCAACGCTGATGAGAGGCGGATGAGACGTCGTC 2580  
Db 6566 CCGGATCTGACGACCGGCTTTCGCTCAACGCTGATGAGAGGCTGATGAGACGTCGTC 6625  
QY 2581 TGTGCTCGATGTCTTACACATGACAGGCGCTGTATCAACGCAATGCGTGGAGAA 2640  
Db 6626 TGTGCTCGATGTCTTACACATGACAGGCGCTGTATCAACGCAATGCGTGGAGAA 6685  
QY 2641 AGCAAGTCCCAATCAACGCTTGAAGCAACTTTTTCGCTGTCGTCACCAACATGCTCA 2700  
Db 6686 ACCAAGTCCCAATCAACGCTTGAAGCAACTTTTTCGCTGTCGTCACCAACATGCTCA 6745  
QY 2701 GCTACACATCCCGCAGCGCAACCGCGGACGAGAAAGGTCACTTTGACAGACTGCA 2760  
Db 6746 GCTACACATCCCGCAGCGCAACCGCGGACGAGAAAGGTCACTTTGACAGACTGCA 6805  
QY 2761 ATCTGAGCGATCACTACAGAGCGTGTTCAGAGAGATGAAAGCGCTCCACAGTT 2820  
Db 6806 GTCTGGAAGACACTACCGGAGCGTGTTCAGAGAGATGAAAGCGCGTCCACAGTT 6865  
QY 2821 AAGCTTAAGTCTTATCAAGAGAGAGCTGCAACCTGACGCGCCCACTTGGGCCAA 2880  
Db 6866 AAGCTTAAGTCTTATCAAGAGAGAGCTGCAACCTGTAAGCTGACGCGCCCACTTGGGCCAA 6925  
|||||

QY 2881 TCTAAATTGGCTATGGGGCAAGAGAGCTCCGAACTTATCCAGCAAGGCAATTACAC 2940  
DB 6926 TCTAAATTGGCTATGGGGCAAGAGAGCTCCGAACTTATCCAGCAAGGCAATTACAC 6985  
QY 2941 ATCCGCTCCGTGTGGAGAGACTGTGGAAAGACCTGAAAACCAATTGACACCAATC 3000  
DB 6986 ATCCGCTCCGTGTGGAGAGACTGTGGAAAGACCTGAAAACCAATTGACACCAATC 7045  
QY 3001 ATGGCAAAAATAGAGTTTCTGTCTCAACCAAGAGAGAGGCGCCAGAGCTTCGC 3060  
DB 7046 ATGGCAAAAATAGAGTTTCTGTCTCAACCAAGAGAGAGGCGCCAGAGCTTCGC 7105  
QY 3061 CTATGCTGTCCAGACTTGGGGGCTCGTGTGCGAGAAAATGGCCCTTATGAGCTG 3120  
DB 7106 CTATGCTGTCCAGACTTGGGGGCTCGTGTGCGAGAAAATGGCCCTTATGAGCTG 7165  
QY 3121 GTCTCAACCTCCCTCAGGCTGTGATGGGCTCCTGTACGGAATTCAGTATTCCTGGA 3180  
DB 7166 GTCTCAACCTCCCTCAGGCTGTGATGGGCTCCTGTACGGAATTCAGTATTCCTGGA 7225  
QY 3181 CAGCGGCTGAGATTCTGTGTGAACGCTTGAAATCAAGAAGACCCCTATGGGCTTTCGA 3240  
DB 7226 CAGCGGCTGAGATTCTGTGTGAACGCTTGAAATCAAGAAGACCCCTATGGGCTTTCGA 7285  
QY 3241 TATGACACCCGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGCTGAGAGAGTCA 3300  
DB 7286 TATGACACCCGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGCTGAGAGAGTCA 7345  
QY 3301 ATTATCAATGTTGTGACTTGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
DB 7346 ATTATCAATGTTGTGACTTGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7405  
QY 3361 CGGCTTTATATCGGGGGTCCCTGACTTATTTCAAAAGGAGAGAGAGAGAGAGAGAG 3420  
DB 7406 CGGCTTTATATCGGGGGTCCCTGACTTATTTCAAAAGGAGAGAGAGAGAGAGAGAG 7465  
QY 3421 TGGCGGCGAGCGGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480  
DB 7466 TGGCGGCGAGCGGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7525  
QY 3481 GCTCTGAGAGCTGTGAG 3540  
DB 7526 GCTCTGAGAGCTGTGAG 7585  
QY 3541 GGGCTTGTGCTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
DB 7586 GAGCTTGTGCTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7645  
QY 3601 TTCAGAGAGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3660  
DB 7646 TTCAGAGAGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 7705  
QY 3661 GACCTGAGAGTGAATCATCATCTCTCCATGATGTGCGTCCGAGAGAGAGAGAGAG 3720  
DB 7706 GACCTGAGAGTGAATCATCATCTCTCCATGATGTGCGTCCGAGAGAGAGAGAGAG 7765  
QY 3721 AAAAGGATATCTACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3749  
DB 7766 AAAAGGATATCTACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 7794

RESULT 13  
AR406043 7989 bp DNA linear PAT 18-DEC-2003  
LOCUS AR406043  
DEFINITION Sequence 10 from patent US 6630343.  
ACCESSION AR406043  
VERSION AR406043.1 GI:40155170  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 7989)

AUTHORS Barteneschlager, R.  
TITLE Hepatitis C virus culture system  
JOURNAL Patent: US 6630343-A 10 07-OCT-2003;  
FEATURES Location/Qualifiers  
source 1. 7989  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 87.8%; Score 3291.4; DB 6; Length 7989;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 1 TGGAGAGGAGCTTTCACAGGCTTACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
DB 3397 TGGAGAGGAGCTTTCACAGGCTTACAGGCTTACAGGAGAGAGAGAGAGAGAGAG 3456  
QY 61 CAGGAG 120  
DB 3457 CAGGAG 3516  
QY 121 CAGGAG 180  
DB 3517 CAGGAG 3576  
QY 181 CAGGAG 240  
DB 3577 CAGGAG 3636  
QY 241 ACAG 300  
DB 3637 ACAG 3696  
QY 301 AG 360  
DB 3697 AG 3756  
QY 361 GAG 420  
DB 3757 GAG 3816  
QY 421 AG 480  
DB 3817 AG 3876  
QY 481 ATGAG 540  
DB 3877 ATGAG 3936  
QY 541 ACAG 600  
DB 3937 ACAG 3996  
QY 601 GAG 660  
DB 3997 GAG 4056  
QY 661 TTGTCACCTCTGCTGAG 720  
DB 4057 TTGTCACCTCTGCTGAG 4116  
QY 721 ACTAG 780  
DB 4117 ACTAG 4176  
QY 781 GCCCACTGCTCCCGAG 840  
DB 4177 GCCCACTGCTCCCGAG 4236  
QY 841 GCTGTTGAG 900  
DB 4237 GCTGTTGAG 4296

QY 901 GGAGTGCAGAGCGCGCTCGTGCCCTTTAAAGTCAAGCGCGCAAAATGCCCTTCCACCGAG 960  
 |||||  
 Db 4297 GGGGTGGCAGGCGCGCTCGTGCCCTTTAAAGTCAAGCGCGCAAAATGCCCTTCCACCGAG 4356  
 QY 961 GACCTGGTTAACTTACTCCCTGCGCATCTCTCTCTGCTGTCCTGGTCCGTGGGGTCTG 1020  
 |||||  
 Db 4357 GACCTGGTTAACTTACTCCCTGCTATCTCTCTCCCTGGCGCCTTAAGTCTGGGGTCTG 4416  
 QY 1021 TGGCGACCGATATCTGGGTGGGCGAGTGGGTCAAGGGGAGGGGCTGTGCAAGTGAAG 1080  
 |||||  
 Db 4417 TGGCGACCGATATCTGGGTGGGCGAGTGGGTCAAGGGGAGGGGCTGTGCAAGTGAAG 4476  
 QY 1081 CGGCTAATAGCGTTTCGCTCGCGGGTAAACATGTTTCCCGCAGCATATGTGCGAGAG 1140  
 |||||  
 Db 4477 CGGCTAATAGCGTTTCGCTCGCGGGTAAACATGTTTCCCGCAGCATATGTGCGCTGAG 4536  
 QY 1141 AGGAGCGCGCAGCAGTGTCACTCAGATCTCTCCGACCTTACTATCAACCACTGTG 1200  
 |||||  
 Db 4537 AGGAGCGCTGCAGCAGTGTCACTCAGATCTCTCTAAGTCTTACATCACTCAGCTGTG 4596  
 QY 1201 AAGAGGCTCCACAGTGAATTAAGAGGACCTGCTCCACCGCCTCGGCTCGTGGCTA 1260  
 |||||  
 Db 4597 AAGAGGCTTCACAGTGAATTAAGAGGACCTGCTCCACCGCAGCTCGGCTCGTGGCTA 4656  
 QY 1261 AGGATGTTTGGGACTGATATGCAAGTTTTGGCTGACTTAAGACCTGCTCAGTCC 1320  
 |||||  
 Db 4657 AGAGATGTTTGGGATGGAATATGCAAGTGTGATGATTTCAAGACCTGCTCAGTCC 4716  
 QY 1321 AAGCTCTGCGCGATTTACCGGAGTCCCTTTTCTCATGCGAAGTGGGTACAAAGGG 1380  
 |||||  
 Db 4717 AAGCTCTGCGCGATTTGCGGAGTCCCTTCTTCTCATGTCAAAGTGGGTACAAAGGG 4776  
 QY 1381 GTCTGGGGGGGAGACGGCATCATGAGAACCAACCGCTCATGTGAGAGCAGATCACCGA 1440  
 |||||  
 Db 4777 GTCTGGGGGGGAGACGGCATCATGAGAACCAACCTGACATGTGAGAGCAGATCACCGA 4836  
 QY 1441 CATGTCAAAAAACGTTTCATGAGGATCGTTGGGCTTAAGACTGTAGTAAATCATGTG 1500  
 |||||  
 Db 4837 CATGTCAAAAAACGTTTCATGAGGATCGTTGGGCTTAAGACTGTAGTAAATCATGTG 4896  
 QY 1501 GGAACATTTCCCATCAACGCAATACCAACGGGCGCTGCAAGCTCTCCGACGGCCAAAC 1560  
 |||||  
 Db 4897 GGAACATTTCCCATCAACGCAATACCAACGGGCGCTGCAAGCTCTCCGACGGCCAAAC 4956  
 QY 1561 TATTTCAAGGGCGCTGTGGCGGGTGTGCTGAGGAGTACGTGAGGTTTACGCGGGTGG 1620  
 |||||  
 Db 4957 TATTTCAAGGGCGCTGTGGCGGGTGTGCTGAGGAGTACGTGAGGTTTACGCGGGTGG 5016  
 QY 1621 GATTTCCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCCCTGCCAGGTTTCA 1680  
 |||||  
 Db 5017 GATTTCCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCCCTGCCAGGTTTCA 5076  
 QY 1681 GCGCCCGAATTTCTTACAGAAATGAGTGGGTGCGCTGCAAGGTACGCTCCGGCGTGC 1740  
 |||||  
 Db 5077 GCGCCCGAATTTCTTACAGAAATGAGTGGGTGCGCTGCAAGGTACGCTCCGACGCTGC 5136  
 QY 1741 AAACCTCTCTACGCGGAGGAGTGCATATTCAGGTCGCGGCTCAACCAATACCTGTGG 1800  
 |||||  
 Db 5137 AAACCTCTCTACGCGGAGGAGTGCATATTCAGGTCGCGGCTCAACCAATACCTGTGG 5196  
 QY 1801 TCGCAGCTCCCATGCGAGCCCGAACCAGATGTAGCACTGCTCACTTCCATGCTCACCGAC 1860  
 |||||  
 Db 5197 TCGCAGCTCCCATGCGAGCCCGAACCAGATGTAGCACTGCTCACTTCCATGCTCACCGAC 5256  
 QY 1861 CCTTCCCAATCAACAGAGAGAGCGGCTTAAAGCGCAGGCTGGCGCAGGGGCTTCCCCCTTC 1920  
 |||||  
 Db 5257 CCTTCCCAATCAACAGAGAGAGCGGCTTAAAGCGCAGGCTGGCGCAGGGGATCTCCCCCTTC 5316  
 QY 1921 TTGGCCAGCTTTTACAGTACGATGTGTGCGGCTTCTCGAAGGAGGACATCATATACC 1980  
 |||||  
 Db 5317 TTGGCCAGCTTTTACAGTACGATGTGTGCGGCTTCTCGAAGGAGGACATCATATACC 5376  
 QY 1981 CAAATATGACTTCCAGAGGCTGACCTCATGAGGCCAACTCTGTGGCGCATGAGATG 2040

Db 5377 CGTCAATGACTCCCGGACGCTGACCTCATGAGGCCAACCTCTGTGGCGGAGAGAG 5436  
 |||||  
 QY 2041 GGGGGGGAATTAACCGGCTGAGTCAAGAAACAAGTATGTAATCTTGACTCTTTTCGAC 2100  
 |||||  
 Db 5437 GGGGGGAATCAACCGGCTGAGTCAAGAAATAAGTATGTAATTTTGGACTCTTTTCGAG 5496  
 QY 2101 CGGCTCCGAGCGAGAGAGATAGCGGGAAGTGTCCGTCCCGGCGAGATCTTCGCGAAA 2160  
 |||||  
 Db 5497 CGGCTCCAGCGGAGAGAGATAGAGAGATCCGTTCCGCGGAGATCTTCGCGAGAG 5556  
 QY 2161 TCCAGAAATTTCCACACAGCATGTCCCGATGGGCAAGCCCGGATTAACAACCTCCGCTG 2220  
 |||||  
 Db 5557 TCCAGAAATTTCCCTCGACAGATGCCAATATGGGACGCGCGGATTTACAACCTTCACATG 5616  
 QY 2221 CTGAGTCTGGAAGGCGCGGACTAAGTCCCTCCAGTGGTACATGGGTGCCACTGCA 2280  
 |||||  
 Db 5617 TTAGAGTCTGGAAGGAGCCGGAATAGTCCCTCCAGTGGTACAGGGTGTCCATTGGCCG 5676  
 QY 2281 CCTACTAAGACCCCTCTCTATACCACTCCACGGAAGAGAGACAGTTGTTTGAAGAA 2340  
 |||||  
 Db 5677 CCTGCAAGGCGCCCTCCGATACCACTCCACGGAAGAGAGACGAGTTGCTCTGACAAA 5736  
 QY 2341 TCAACGTGTCTTTCGCTCGGAGCTTGGCCAAAGGCTTTTGGTATGCTCCGAACCG 2400  
 |||||  
 Db 5737 TCTACCGTGTCTTTCGCTCGGAGCTCGCACAACCAACCTTGGCAGCTCCGATG 5796  
 QY 2401 TCGGCGGTGACAGCGGCAACCGCCCTCTGACCAACCTTCCGACGCGGCA 2460  
 |||||  
 Db 5797 TCGGCGGTGACAGCGGCAACCGCCCTCTGACCAACCTTCCGACGCGGCA 5856  
 QY 2461 GCAGATCTGAAGTGAAGTGTATTCCTCATGCCCCCTTTGAGGGGAGCCGCGGAGAC 2520  
 |||||  
 Db 5857 GGGGATCCGAGGTTGAAGTGTATCTCTCATGCCCCCTTTGAGGGGAGCCGCGGAGAT 5916  
 QY 2521 CCGATCTGAGGACGGGTCTTGTGTACCGTGAAGTGAAGGCGGCTGAAGACCTCTGTC 2580  
 |||||  
 Db 5917 CCGATCTGAGGACGGGTCTTGTGTACCGTGAAGTGAAGGCGGCTGAAGACCTCTGTC 5976  
 QY 2581 TGCTGCTGATGTCTTACATGAGCAGGCGCTGTATCAGCAGCATGCGCTGCGAGAGAA 2640  
 |||||  
 Db 5977 TGCTGCTGATGTCTTACATGAGCAGGCGCTGTATCAGCAGCATGCGCTGCGAGAGAA 6036  
 QY 2641 AGCAAGCTGCCATCAACCGCTTGAAGCAACTTGTGCTGCTGCTACCAACAATGATCTAC 2700  
 |||||  
 Db 6037 ACCAAGCTGCCATCAATGATGATGACCACTGTTGCTCGTACCAACAATGATCTAC 6096  
 QY 2701 GCTTACCAATCCCGCAGCGCAAGCCAGCGGCAAGAGAGTCACTTTTGAACAATGCTGCA 2760  
 |||||  
 Db 6097 GCTTACCAATCCCGCAGCGCAAGCCAGCGGCAAGAGAGTCACTTTTGAACAATGCTGCA 6156  
 QY 2761 ATCTGGAAGATACCTACAGAGAGCGGTCAAGAGATGAGAGGCGAAGGCGTCAAGTT 2820  
 |||||  
 Db 6157 GTCTGGAAGACACTTACCGGAGCGTGTCTGAGAGATGAGAGGCGAAGGCGTCAAGTT 6216  
 QY 2821 AAGGCTAAGCTTCTATCAGTGAAGAGGCTGCAAGCTGACCGCCCAATTCGCGCAAA 2880  
 |||||  
 Db 6217 AAGGCTAAGCTTCTATCAGTGAAGAGGCTGCAAGCTGCAAGCTGACCGCCCAATTCGCGCAAG 6276  
 QY 2881 TCTAAATTTGGCTATGGGCAAGAGCGTCCGGAACCTTATCCAGCAAGGCGGTTAACAG 2940  
 |||||  
 Db 6277 TCTAAATTTGGCTATGGGCAAGAGCGTCCGGAACCTTATCCAGCAAGGCGGTTAACAG 6336  
 QY 2941 ATCCGCTCGTGTGGAGAGACTTGTGGAAGACACTGGAACCAATTTGACACCAATTC 3000  
 |||||  
 Db 6337 ATCCGCTCGTGTGGAGAGACTTGTGGAAGACACTGGAACCAATTTGACACCAATTC 6396  
 QY 3001 ATGGCAAAAAATAGAGTTTTCGCTGCTCAACAGAGAGAGAGGCGCAAGCCAGCTCGC 3060  
 |||||  
 Db 6397 ATGGCAAAAAATAGAGTTTTCGCTGCTCAACAGAGAGAGAGGCGCAAGCCAGCTCGC 6456  
 QY 3061 CTTATGCTGTTCCAGACTTTGGGGGTCCGTGTGCGAAGAAATGGCCCTCTATGACGTG 3120  
 |||||



QY	1081	GGCGTGAATGAGGTTGGCTCCGGGGGTAAACAATGTTTCCCAAGCACTAATGTCAGAG	1140
Db	4477	CGCGTAATGAGCGTTGGCTTCGGGGGTAAACAACGCTCCCTCCACGCACTAATGTCCTGAG	4536
QY	1141	AGCGACGGCCGAGCAAGTGTCACTAGATGCTCTCCGACCTTAACTATCAACCACTGTG	1200
Db	4537	AGCGACGCTGAGCAAGCTGTCACTAGATCTCTCTAATCTTTACATCACTCACTGCTG	4596
QY	1201	AAGAGGCTTCACAGATGATTTAACAGAGACTGTCCACAGCCCTGTCCGCTCGTGCCTA	1260
Db	4597	AAGAGGCTTCACAGATGATTCACAGAGACTGTCCACAGCCATGCTCCGCTCGTGCCTA	4656
QY	1261	AGGATGTTTTGGGACTGTGATATGCAACTGTTTTGGGTGACTTCAAGACTGTGCTCAATCC	1320
Db	4657	AGAGATGTTTTGGGATTTGGATATGCAACGGTGTGTACTGATTTCAAGACTGTGCTCAATCC	4716
QY	1321	AAGCTCTGCGCGCAATTAACCGGAGATGCCCTTTTTCATGTCGAACGATGGGTACAAAGGG	1380
Db	4717	AAGCTCTGCGCGCAATTTGCGGAGATGCCCTTTTTCATGATCAACGATGGGTACAAAGGA	4776
QY	1381	GTCTGGCGGGGAGACGGGCAATGACAGCAACCTGTCTATGTGAGACAGATCAACCGGA	1440
Db	4777	GTCTGGCGGGGAGACGGGCAATGACAAACCACTGTCCCATGTGGAGACAGATCAACCGGA	4836
QY	1441	CATGTCAAAAAAGTTTCCATGATAGATTCGTTGGGCTTAAGACTCTGATATGTAATGTGGCAT	1500
Db	4837	CATGTGAAAAAGTTTCCATGATAGGATTCGTGGGCTTAAAGACTCTGATATMACCTGTGCAT	4896
QY	1501	GGAACTATTCATCAACGCAATPACACACAGGGGCCCTGTACAGCCCTCCCGGAGGCAAC	1560
Db	4897	GGAACTATTCATTCATTAACGCGTACACACAGGGGCCCTGTACAGCCCTCCCGGAGGCAAT	4956
QY	1561	TATTCAGGGCGCTGTGCGGGGTGCTCTGAGAGTACGTGAGGTTACCGGGGTGGG	1620
Db	4957	TATTCAGGGCGCTGTGCGGGGTGCTCTGAGAGTACGTGAGGTTACCGGGGTGGG	5016
QY	1621	GATTTCCTACTAGTACAGACGATGCCACTGACAACTGTAATGCGCTGCAGGTTCCA	1680
Db	5017	GATTTCCTACTAGTACAGGAGATGCCACTGACAACTGTAATGTCCTGTCAAGTTCCG	5076
QY	1681	GCCCCGAAATCTTACAGAAAGTGAATGGGGGTGGGCTGCACAGGTACGCTCGGCGGTGC	1740
Db	5077	GCCCCGAAATCTTACAGAAAGTGAATGGGGGTGGGCTGCACAGGTACGCTGCAGCGTGC	5136
QY	1741	AAACCTCTCTTACGGGAGAGGTCACTATTCAGAGTCCGGGCTCAACCAATACCTGTTGGG	1800
Db	5137	AAACCTCTCTTACGGGAGAGGTCACTTCTGTGTGGGCTCAATCAATACCTGTTGGG	5196
QY	1801	TCGCAGCTTCCATGCGAGCCGAAACCGGATGTAGCATGTCTCACTTCATGCTCACCGAC	1860
Db	5197	TCACAGCTTCCATGCGAGCCCGAAACCGGACGTAGCATGTCTCACTTCATGCTCACCGAC	5256
QY	1861	CCCTCCCACTACAGCAGAGACGGCTAAAGGCGAGCTGGGCAAGGGGTCTTCCCCCTCC	1920
Db	5257	CCCTCCCACTTACGGCGGAGACGGCTTAAAGGCTGAGCTGGGCAAGGGATCTTCCCCCTCC	5316
QY	1921	TTGGCAGGCTTCTTCACTTACAGCAAGTGTCTGCGCTTCTCTGAAAGGCACTATCACTTACC	1980
Db	5317	TTGGCAGGCTATCACTTACAGCAAGCTGTCTGCGCTTCTCTTGAAGGCACTATGACTACC	5376
QY	1981	CAAAATGACTTCCAGACGCTGACTCTCATGAGGCAACCTTGTGGCGGCAATGAGATG	2040
Db	5377	CGTCACTGACTCCCGGAGCGCTGACTCTCATGAGGCAACCTTGTGGCGGCAATGAGATG	5436
QY	2041	GGCGGGGACATTAACCCGCTGTGAGTGCACAGAAACAAGGTATGTAATCTTGAACTCTTTGAC	2100
Db	5437	GGCGGGGACATCAACCCGCTGTGAGTGCACAAAAATAAGTATGTAATTTTGAACTCTTTGAG	5496
QY	2101	CGCGCTCCAGCGAGAGATGAGGAGGAAAGTGTCCGTCCCGGCGGAGATCTGCGGAAA	2160
Db	5497	CGCGCTCAAGCGAGAGATGAGGAGGAAAGTATTCGTTCCGCGGAGATCTTCCGAGG	5556
QY	2161	TCCAAGAAATTCACACAGCGATGCTCCCATGAGGCAACCCCGGATTTACACCTTCGCTG	2220

Db	5557	TCGAGAAATTCCTCGAGCAAGTCCCATATGGGAGCGCCGGAATTCACCTTCACATG	5616
Oy	2221	CTGAGTCTCTGAAAGGCCCGGACCTACGTCCCTCAGTGTACATGGGTGCCCATGCA	2280
Db	5617	TTAGAGTCTCTGGAAGGACCCGGGACTAGTCCCTCAGTGTACACGGGTGTCCATTGGC	5676
Oy	2281	CTTACTAAGACCCCTCTTATACACTCTCAACGAGAAAGAGACAGTTGTTCTGACGAA	2340
Db	5677	CCTGCGCAAGGCCCTCCGATACACCTTCACGAGAGAAAGAGCGTTGTCTCTGAGAA	5736
Oy	2341	TCACCGGTCTTCGCGCTGGCGGAGCTTGGCCAAAGGTTTCGGTAGTCCGAACG	2400
Db	5737	TCATCCGTCTTCCTGCTTGTGGCGAGTCCGCAAAAGACTTTCGGAGCTTCGAATCG	5796
Oy	2401	TCGCGCGTTCGACAGCGGACCGGCAACCGCCCTCTGACCAACCTCCGACGAGCGGGA	2460
Db	5797	TCGCGCGTTCGACAGCGGACCGGCAACCGCCCTCTGACCAACCTCCGACGAGCGGAC	5856
Oy	2461	GCAAGATCTGACGTTGAGTGTGATTCCTTCATGCCCCCTTTGAGGGGAGCCGGGGAC	2520
Db	5857	GCGGAGATCCGACGTTGAGTGTGATTCCTTCATGCCCCCTTTGAGGGGAGCCGGGGAT	5916
Oy	2521	CCGATCTCAACGACGGGTCTTGTGTCTACCGTGAGTACGAGGCGCGGTGAGGAGTGTGC	2580
Db	5917	CCGATCTCAACGACGGGTCTTGTGTCTACCGTGAGTACGAGGCGCGGTGAGGAGTGTGC	5976
Oy	2581	TGCTGCTGATGTCTCTACACATGACAGGCGCTCTGTATCAGCCATCGCTCCGAGGAA	2640
Db	5977	TGCTGCTGATGTCTCTACACATGACAGGCGCTCTGTATCAGCCATCGCTCCGAGGAA	6036
Oy	2641	AGCAAGCTGCCCATCAACGCGTTGAGCACTCTTGTCTGCTGCACCAACATGATGTAC	2700
Db	6037	ACCAAGCTGCCCATCAACGCGTTGAGCACTCTTGTCTGCTGCACCAACATGATGTAC	6096
Oy	2701	GCTACCAATCCCGGAGGCGAAGCCAGGGGAGAAAGTCACTTTGACAGACTCGAA	2760
Db	6097	GCTACCAATCCCGGAGGCGAAGCCAGGGGAGAAAGTCACTTTGACAGACTCGAG	6156
Oy	2761	ATCCGAGACATCACTACACAGACGTCCTCAAGAGATGAAAGGCGAAGCGTCCACAGTT	2820
Db	6157	GTCCTGAGACACACTACCGGGAGCTGTCTCAAGAGATGAAAGGCGAAGCGTCCACAGTT	6216
Oy	2821	AAGGTAAGCTTCTATCAGTAGAGAAAGCTGCAAGCTGACGCCCCACATTTGGGCCAA	2880
Db	6217	AAGGTAAGCTTCTATCAGTAGAGAAAGCTGCAAGCTGACGCCCCACATTTGGGCCAA	6276
Oy	2881	TCTAATTTGGCTATGGGGCAAAGAGCGTCCGAACTATCAGCAAGGCCATTACAC	2940
Db	6277	TCTAATTTGGCTATGGGGCAAAGAGCGTCCGAACTATCAGCAAGGCCATTACAC	6336
Oy	2941	ATCCGCTCCGAGTGGGAGACCTTGTGGAGAACCTGAAACACCAATTGACACCAATC	3000
Db	6337	ATCCGCTCCGAGTGGGAGACCTTGTGGAGAACCTGAAACACCAATTGACACCAATC	6396
Oy	3001	ATGCGAAAAATGAGGTTTCTGGCTCCACACAGAGAGAGGCGCGAAGCCAGCTGCG	3060
Db	6397	ATGCGAAAAATGAGGTTTCTGGCTCCACACAGAGAGAGGCGCGAAGCCAGCTGCG	6456
Oy	3061	CTTATCGTTTCCAGACTTGGGGGTCCGTGTGTGCGAAGAAATGGCCCTTATGACGTG	3120
Db	6457	CTTATCGTTTCCAGACTTGGGGGTCCGTGTGTGTGCGAAGAAATGGCCCTTATGACGTG	6516
Oy	3121	GTCACACCCCTCCAGCGCTGTATGGGCTCCGTGACGATTCACGATTCCTCTGGA	3180
Db	6517	GTCACACCCCTCCAGCGCTGTATGGGCTCCGTGACGATTCACGATTCCTCTGGA	6576
Oy	3181	CAGCGGCTGAGTTCTGTGTGAACGCTCGAATAACAAAGAACCCCTATGGCTTTGCA	3240
Db	6577	CAGCGGCTGAGTTCTGTGTGAATGCTCGAATAACAAAGAACCCCTATGGCTTTGCA	6636
Oy	3241	TATGACACCGCTGTTTGTACTCAACAGTCACTGAGATGACATCCGTGAGAGAGTCA	3300



Db 6637 TATGACACCCGCTGTTTGAAGCTCAACGCTGACGATGATGATCATCCGTTTGAAGAGTCA 6696

Qy 3301 ATTATCAATGTTGTGACTTGCGCCCGGAGGACGAGCGCATTAAGTGGCTCAGAG 3360

Db 6697 ATCTACCAATGTTGTGCTTGCGCCCGGAGGACGAGCGCATTAAGTGGCTCAGAG 6756

Qy 3361 CGGCTTATATCGGGGGTCCCTGACTAATTTCAAAAAGGAGAACTGGCTATCGCGG 3420

Db 6757 CGGCTTATACATCGGGGGGCCCCCTGACTAATTTCTTAAAGGAGAACTGGCTATCGCGG 6816

Qy 3421 TGGCGCGGAGCGCGCTGCTGACGACTAGCTGCGGTAAATACCTCACAATGTTACTTGAAG 3480

Db 6817 TGGCGCGGAGCGCGTGTACTGACGACGAGCTGCGGTAAATACCTCACAATGTTACTTGAAG 6876

Qy 3481 GCGCTGACGCTGTGAGCTGCAAAAGCTCCAGACATGACGATGCTCGTGTGCGGAGAC 3540

Db 6877 GCGGCTGCGGCTGCTGAGCTGCGAAAGCTCCAGAGCTGACGATGCTGTGATGCGGAGAC 6936

Qy 3541 GGCCTTGTCTGTTATCTGTGAGAGCGCGGAAACCGAGAGACGCGCGGAGCTTACGAGTC 3600

Db 6937 GACTTGTCTGTTATCTGTGAAAGCGCGGAGACCGAAGAGAGCGGAGCTTACGAGTC 6996

Qy 3601 TTCACGAGGCTATGACTAGTACTGCTGCCCCCGGGAGACCGCCCAACGAGAAATAC 3660

Db 6997 TTCACGAGGCTATGACTAGTACTGCTGCCCCCGGGAGACCGCCCAACGAGAAATAC 7056

Qy 3661 GACTTGTGATGATTAATATATGCTCTCCATGTCGCTGCGGACGATGATCTGCGC 3720

Db 7057 GACTTGTGATGATTAATATATGCTCTCCATGTCGCTGCGGACGATGATCTGCGC 7116

Qy 3721 AAAAGGTATATACCTCAGCCCGTGAACC 3749

Db 7117 AAAAGGTATATCTCAACCCGTGAACC 7145

RESULT 15

SSE242652

LOCUS Hepatitis C virus replicon 1377/NS3-3'UTR. linear SYN 11-MAY-2000

DEFINITION

ACCESSION AJ242652.1 GI:5441834

KEYWORDS core-neo gene; NS2 gene; NS2 proteinase; NS3 gene; NS3 proteinase/helicase; NS3/4A proteinase cofactor; NS4A gene; NS4B gene; NS4B protein; NS5a gene; NS5a phosphoprotein; NS5b gene; NS5b RNA dependant RNA polymerase; polypotein.

SOURCE Hepatitis C virus replicon 1377/NS3-3'UTR

ORGANISM Hepatitis C virus replicon 1377/NS3-3'UTR

REFERENCE 1 other sequences; artificial sequences; vectors.

AUTHORS Lohmann, V., Korner, F., Koch, J., Heitan, U., Theilmann, L. and Bartenschlager, R.

TITLE Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line

JOURNAL Science 285 (5424), 110-113 (1999)

MEDLINE 99322193

PUBMED 10390360

REFERENCE 2 (bases 1 to 7989)

AUTHORS Bartenschlager, R.

TITLE Direct Submision

JOURNAL Submitted (26-MAY-1999) Bartenschlager R., Institute for Virology, Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67, 55131 Mainz, GERMANY

FEATURES

source

1. .7989 Location/Qualifiers

/organism="Hepatitis C virus replicon 1377/NS3-3'UTR"

/mol\_type="other RNA"

/db\_xref="taxon:95363"

/focus

/note="Neomycin-selectable bicistronic subgenomic hepatitis C virus replicon"

join(1..341,1801..7758)

/organism="Hepatitis C virus"

/mol\_type="other RNA"

/strain="1b"

/isolate="Cont"

/db\_xref="taxon:11103"

/cissue\_type="liver"

/cissue\_type="liver"

/organism="Encephalomyocarditis virus"

/mol\_type="other RNA"

/db\_xref="taxon:12104"

1..341

/evidence=experimental

342..1181

/gene="core-neo"

342..1181

/function="selectable marker"

/codon\_start=1

/product="core-neo fusion protein"

/protein\_id="CAB46912.1"

/db\_xref="GI:5441835"

/translation="MSTNPKPQKTKGRAMIEDQGHAGSPAAWRLFGYMAOQTI

GCSDAVPRHLSAGRPVLTGKIDLSGALNEODEARLSMLATGVPICAVLVITVEA

GRDWLLGEVPGDQLSSHAPAEKVSINADARRLLHTIDPATCPDPOKRIEAK

TRMEAGLVQDDLDDEHOGIAPAELEFARLKARMPDEDLVTHGDCLPINWENRF

SGFIDCGRLGVADRYODIALATRDIAEIEIGEWADRFVLVYGIAPDSQRIAFYRLD

EPF"

1190..1800

/note="internal ribosome entry site (IRES)"

1801..7758

/product="non-structural polyprotein"

/codon\_start=1

/protein\_id="CAB46913.1"

/db\_xref="GI:5441836"

/translation="MAPITRYSQTRGILGCIITSITGRDRNOVEGVYVSTRTGS

LATCVAGCVTHVHGASRTLAGPKPIITMTNVQDILVGMAPGANSVITRTGS

SDLYLVTRADVIIVRRGDSRGLSPRVSLKSSGGPLCPSGHAGVITRAVCA

TRGVAVADVFPVPESMETTRSPVFTDNGSPVAVPOTFOVAHIAATPGSKTKRAVC

YAAOGKVLVLPNSVATLGFVAMSKAGIDINRTGATITTTGAPITVSTGKFLA

DGCGSGAVDIITCDECHSTDTITLIGIVLDOARTAGRLVAVLAPGSGVTPH

PIREYVALSTGEIIPFYKAIPIETIKGRHLIFCHSKKKCDLAKLGLGIANAVY

YRGLDVSIPTSGDVIIVATDALMTGTGPDSVIDICNTCTOTVTPDSDPTTIETT

TYPDVAVSRQRRGTGRGMGIYRPVTPERSGMPDSSVLCECYDAGCAWELTPA

ETSVRAVLTNPGLPVCODHLEFMESVFTGLPHIDAHPLSQTKQAGDNPVYALQA

TYCARADAPPSPMDOMKCLIRKPLTHTPILYLVGAVONVTTHTPTKYMAG

SALREYVTTWVGVVLAALAAVCLTSSVSIVGIIISGKALIIDBEVLYRDE

MECASHLPIYEGMQLAQFKOKAIGLIQTRAKQAAAPVYESKRTLEAWAGM

WNFTSGIYIAGISTIPGNPALSLAFLASTISPLTGHLLFNLLGEMVAQLAP

SAASAVGAGIAGAAVGSIGKVLVDIILAGAGAVGALVAFKNGSGMSTEDLVN

LLPAIISPGALVGVVCAALIRRHVPGGGAQVMNRLIAFASRGHNSPTHTVPESD

AAARVQIISLITITQLIKRLHOWINEDSTPCGSGMLRDVMDMICTVLIDFTWLOS

KILPRLPVYPPSCORGVGVWRGDIINOTTCGCAQITGVGHGSMRLVIGPRTGNT

WCHTPIINAVTTPGCTPSPAVNSRALNRYVAEYEVITRVGDPHVYTGCTDYNKCP

COVPAPEPTEVDGVLRLHYAPACKPLAEVEVITGLVQVLSQDPECEPDDVAVL

TSMLTDPHSHTAETAKRLARLGSPLSLASSASQSLAPSLKATCTTRHSDPPDLLEA

NILMRQEMGNITRVESENKVVILDFEELQAEEDREVSVAPEILRRSKFRPAPPI

MAAPDYNPLPLESMKDPDYVPVVGCCPLPPAKAPPIPPRRKRTVLSSTVSALA

ELATKTRGSSSAVDGCTATASPPDSDGAGASVESYSSMPLEGGEGDDLSDG

SNSTVSEASSEDVCCSMSTYTGALITTCABETKLPIINALNSLIRHNLVYATTS

RSASLROKVTEDRLQVLDHYRDVLKEMKAKASTYKALKLSVEEKALPIPTASAK

FGAGADVRLNLSKAVNHLRSWKDLLEDTETPIDTTINAKNEFCVQEKGRKAR

LIVFPLGVAVCEKALVDVSTLPQAAVSGVSPGQREYFVLMVAKKQNG

FAVDNRCPDSTENDIRVESIYOCCLDAPBARQAIRSTELTYGGPLTNSKGC

GVRRCASGVLTTCGNTLTLYKAAARAKLQDCTVMAVCSDDLVVICSAGQDM

EASLRATEAMTRYSAPPDGPPKPEYDELITSCSSNVYAHDAKSKRYVILTRDPT

PLARAWETARHTPVNSWGNITIMVAPITMARINLMTHTFSLIAGEOLKALDCIY

GACYSIEPLDLQIQRHLGLSAFSLHSSPGKINVASCLRLDGVPPRLVWHRARS

VRARLSQGRATACGYLFENNAVTRKLTLPPIPASQDLDSMFWAGVSGDIYHSL

SRAPRPMFWMCILLISVGVYILPBR"

1801..3596

/gene="NS3"

/product="NS3 proteinase/helicase"

3697..3858

/gene="NS4A"

/product="NS3/4A proteinase cofactor"

mat\_peptide

mat\_peptide

```
mat_peptide      3859..4641
                  /gene="NS4B"
                  /product="NS4B protein"
mat_peptide      4642..5982
                  /gene="NSA"
                  /product="NSA phosphoprotein"
mat_peptide      5983..7755
                  /gene="NS5B"
                  /product="NS5B RNA dependant RNA polymerase"
gene             1801..3696
                  /gene="NS3"
                  /gene="NS3"
gene             3697..3858
                  /gene="NS4A"
                  /gene="NS4A"
gene             3859..4641
                  /gene="NS4B"
                  /gene="NS4B"
gene             4642..5982
                  /gene="NSA"
                  /gene="NSA"
gene             5983..7755
                  /gene="NS5B"
                  /gene="NS5B"
gene             7759..7989
ORIGIN
3'UTR
Query Match      87.8%; Score 3291.4; DB 12; Length 7989;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY      1 TGGAGAGGCGCTCTTCAAGAGCTCAACCCAGTGATGCCACTTCTGTCCTCCAAACAAAG 60
DB      3397 TGGAGAGCGCTCTTACAGGCTCAACCAATAGACGCCCAATTTCTGTCCTCAAGACTAG 3456

QY      61 CAGGACAGAGACAACCTTCCCTACCTGGTGCGTACAGAGCTACTGTGTGCGGTAGGGCC 120
DB      3457 CAGGACAGAGACAACCTTCCCTACCTGGTACATACAGAGCTACGGTGTGCGCAGAGGCT 3516

QY      121 CAGGCCCACTCTCATATATGGGATCAATGTGGAAGTGTCTCATACGGCTAAAGCCTAAT 180
DB      3517 CAGGCTCACTCTCATGTGTGGACCAATGTGGAAGTGTCTCATACGGCTAAAGCCTAAG 3576

QY      181 CTGCGCGGGGCAACACCTTGTGTATAGGCTGGAGACCGTCCAAACGAGATCAACCTCT 240
DB      3577 CTGCGCGGGGCAACGCCCTGTGTATAGGCTGGAGACCGTCCAAACGAGATTAATCAAC 3636

QY      241 ACAACCCCAATTAACCAATTCATCATGAGCATGTGTCAAGCCGACCTGAGAGTGTGACG 300
DB      3637 ACAACCCCAATTAACCAATTCATCATGAGCATGTGTGAGCTGACCTGAGAGTGTGACG 3696

QY      301 AGCACTGGGTGTGTGTGGGCGGGGTCTCTTCAAGCTCTGGTGTGTATTTGCTTGAACA 360
DB      3697 AGCACTGGGTGTGTGTGGGCGGGGTCTCTTCAAGCTCTGGTGTGTATTTGCTTGAACA 3756

QY      361 GGCAGCGTGTGATTTGTGGTAGGATCATCTGTCCGGGCGGCGGCTATTTGTTCCCGAC 420
DB      3757 GGCAGCGTGTGATTTGTGGTAGGATCATCTGTCCGGGCGGCGGCTATTTGTTCCCGAC 480

QY      421 AGGGAAGTCTCTTACAGAGTTGATGAGATGGAAGAGTGTGCGGTGCACTTCCCTTAC 480
DB      3817 AGGGAAGTCTCTTACCGGAGTTGATGAGATGGAAGAGTGTGCGGTGCACTTCCCTTAC 3876

QY      481 ATCGAGCAGGGAATGTCACTGTGCGGAGAGATTCAAGCAAAAAGCGCTCGGTTGCTGAC 540
DB      3877 ATCGAGCAGGGAATGTCACTGTGCGGAGCAATTTCAAAAGGAAGCGGTTGCTGCA 3936

QY      541 ACAGCCCAAGCAAGGAGGAGCGGCTGTCCGTTGGAGTCAAGTGGCGAGCCCTT 600
DB      3937 ACAGCCCAAGCAAGGAGGAGCGGCTGTCCGTTGGAGTCAAGTGGCGAGCCCTT 3996

QY      601 GAGACTTTTGGGCGAAACCATGTGGAATTTATCAAGCGGATACATTAAGCAGC 660
DB      3997 GAGACTTTTGGGCGAAACCATGTGGAATTTATCAAGCGGATACATTAAGCAGC 4056

QY      661 TTGTTCATCTGCTGGGAATCCGCGATTTGCATCACTGATGGGTTCAACAGCTCTGTCT 720
DB      4057 TTGTTCATCTGCTGGGAATCCGCGATTTGCATCACTGATGGGATTTCAACAGCTCTTATC 4116
```

```
QY      721 ACTAGCCGCTCAACCAACCAATTCATCCTCTGCTTAAACATCTCGGGGAGATGGTAGCC 780
DB      4117 ACCAGCCGCTCAACCAACCAATTCATCCTCTGCTTAAACATCTCGGGGAGATGGTAGCC 4176

QY      781 GCCCAACTGCTTCCCCCAAGTGTGTGAGCTTTGTGTGTGGGCGGCAATTTGCTGTGG 840
DB      4177 GCCCAACTGCTTCCCCCAAGTGTGTGAGCTTTGTGTGTGGGCGGCAATTTGCTGTGG 4236

QY      841 GCTGTGGAGCATAGAGCTTGGGAGAGTGTGTGAGATCATCTGTGGCGGCTATAGAGCA 900
DB      4237 GCTGTGGAGCATAGAGCTTGGGAGAGTGTGTGAGATCATCTGTGGCGGCTATAGAGCA 4296

QY      901 GGAAGTGGAGCGCGCTGTGTGCTTTAAGTATAGAGCGGGAATGCCCTTCCACCGAG 960
DB      4297 GGAAGTGGAGCGCGCTGTGTGCTTTAAGTATAGAGCGGGAATGCCCTTCCACCGAG 4356

QY      961 GACTGTGTTAATCTATCTCCCTGCAATCTCTCTCTCGTGTGCGCTGGTCCGCGGGTCTG 1020
DB      4357 GACTGTGTTAATCTATCTCCCTGCAATCTCTCTCTCGTGTGCGCTGGTCCGCGGGTCTG 4416

QY      1021 TGGCAGCGATATCTGGGTGGGACGTCAGGAGTCCAGAGGAGGAGGAGCTGTGAGTGAAC 1080
DB      4417 TGGCAGCGATATCTGGGTGGGACGTCAGGAGTCCAGAGGAGGAGGAGCTGTGAGTGAAC 4476

QY      1081 CGGCTGATAGCGTTGCGCTCGCGGGTAAACATGTTTCCCAACGCACTATGTGCCAGAG 1140
DB      4477 CGGCTGATAGCGTTGCGCTCGCGGGTAAACAGTCTTCCCAACGCACTATGTGCTGAG 4536

QY      1141 AGGAGCGCGGACAGAGTGTCACTCAATGATCTCTCCGACTTATCAACCACTGTGG 1200
DB      4537 AGGAGCGTGAACAGAGTGTCACTCAATGATCTCTCTCAATGATCTCACTCACTGCTG 4596

QY      1201 AAGAGGCTCAACAGATGATTAACGAGACCTGCTCCAGCCCTGCTCCGGCTGATGGCTA 1260
DB      4597 AAGAGGCTCAACAGATGATTAACGAGACCTGCTCCAGCCCTGCTCCGGCTGATGGCTA 4656

QY      1261 AAGGATGTTTGGAGCTGATATGACAGATTTTGGCTGACTTCAAGCCTGCTCAAGTCC 1320
DB      4657 AAGGATGTTTGGAGCTGATATGACAGATTTTGGCTGACTTCAAGCCTGCTCAAGTCC 4716

QY      1321 AAGCTCTGCGGAGTATCCGGAGTCCCTTTTTCATGCCAACAGTGGGTCAAGGGG 1380
DB      4717 AAGCTCTGCGGAGTATCCGGAGTCCCTTTTTCATGCCAACAGTGGGTCAAGGGG 4776

QY      1381 GTCTGGGAGGAGACGCGCATATGACAGCAACCTGTCTCATGTGTGAGCAAGATCACCGGA 1440
DB      4777 GTCTGGGAGGAGACGCGCATATGACAGCAACCTGTCTCATGTGTGAGCAAGATCACCGGA 4836

QY      1441 CATGTCAAAAACGCTTCAATGAGATCGTTGGGCTTAAGACCTGTAGTAAATGTGGCAT 1500
DB      4837 CATGTCAAAAACGCTTCAATGAGATCGTTGGGCTTAAGACCTGTAGTAAATGTGGCAT 4896

QY      1501 GGAACATTTCCCATCAACGATATACCAAGGAGGCGCTTGTGACAGCGCTCCCAAGGCCA 1560
DB      4897 GGAACATTTCCCATCAACGATATACCAAGGAGGCGCTTGTGACAGCGCTCCCAAGGCCA 4956

QY      1561 TATTTCAGGCGGTGTGGGCGGAGTGTGCTGTGAGAGTAAAGTGTGAGGTTACGCGGTTGG 1620
DB      4957 TATTTCAGGCGGTGTGGGCGGAGTGTGCTGTGAGAGTAAAGTGTGAGGTTACGCGGTTGG 5016

QY      1621 GATTTCCACTACGTGACGAGATGACCACTGACAAAGTAAATGCCGTGACAGTTTCA 1680
DB      5017 GATTTCCACTACGTGACGAGATGACCACTGACAAAGTAAATGCCGTGACAGTTTCA 5076

QY      1681 GCCCCGGAATTTCTTACAGAGTGAATGGGTGTGCGTCAAGAGTACGTTCCGCGGTGTC 1740
DB      5077 GCCCCGGAATTTCTTACAGAGTGAATGGGTGTGCGTCAAGAGTACGTTCCGCGGTGTC 5136

QY      1741 AAACCTCTCTAGGGAGGAGGTCAATTCAGTGGGCTCAACCAATCACTGTGTTGG 1800
DB      5137 AAACCTCTCTAGGGAGGAGGTCAATTCAGTGGGCTCAATTCAGTGTGTTGG 5196
```



QY	1801	TCGGAGCTCCCATGGGAGCCCGAACCAGATGTAAGAGTCTCACTTCCATGCTCAACGAC	1860
Db	5197	TCACAGCTCCCATGGGAGCCCGAACCAGATGTAAGAGTCTCACTTCCATGCTCAACGAC	5256
QY	1861	CCCTTCCACATACAGCAGAGACGGCTTAAGGCGCAGGCTGGGCGAGGGGGCTTCCCTCC	1930
Db	5257	CCCTTCCACATTTAAGGGGGAGACGGCTTAAGGCGTGAAGCTGGGCGAGGGGATCTTCCCTCC	5316
QY	1921	TTGGCCAGCTCTTCAGCTAGCCAGTTGTCGTGGCTTCTTCGAAGGCGACATACATTACC	1980
Db	5317	TTGGCCAGCTCATCAGCTAGCCAGCTGTCGTGGCTTCTTCGAAGGCGACATGACCTACC	5376
QY	1991	CAAAATGACTTCCCGAGAGCTTGACCTCATCGAGGCAACCTTCGTGGCGGCATGAGATG	2040
Db	5377	CGTACATGACTCCCGAGAGCTTGACCTCATCGAGGCAACCTTCGTGGCGGCAGAGATG	5436
QY	2041	GGCGGGGACATTTACCGCGGTGAAGTCAGAGAAACAAGTAGTATCTTGGACTCTTTGCAC	2100
Db	5437	GGCGGGAACTACACCGCGGTGAAGTCAGAGAAATAGTAGTATTTTGGACTCTTTGCAG	5486
QY	2101	CCGCTCCGAGCGGAGAGAGATGAGCGGGAGTGCCTCCGCGCGAGATCTTCGAGAA	2160
Db	5497	CCGCTCCAAGCGGAGAGAGATGAGGAGAAATGATCCGTTCCGCGCGAGATCTTCGAGAG	5556
QY	2151	TTCCAGAAATTTCCCAACGAGCATGCGCGCATGGGACGCGCGGATTAACAACCTTCGCTG	2220
Db	5557	TTCCAGAAATTTCCCTCGAGCGCATGCGCATATGGGACGCGCGGATTAACAACCTTCAC	5616
QY	2221	CTGAGTCTTGAGAGGCGCGCGGACCTACGCTCCAGAGTATACATGGGTGCGCACTGCA	2280
Db	5617	TTAGAGTCTTGAGAGGACCTCGGACTACGCTCCCTCGAGTATACAGGATGCTCATTTGCC	5676
QY	2281	CCTACTAAGACCCCTCTATACCACTTCCAGAGGAAAGAGACAGTTGTTCTGACAGAA	2340
Db	5677	CCTGCAAGGCGCCCTCCGATACACTCCAGAGGAAAGAGACGATTTGTCCTGTCAGAA	5736
QY	2341	TTCAACCGTGTCTTTCGCCCTGGCGGAGCTTGCACAAAGGCTTTGGTAGCTCGGAACCG	2400
Db	5737	TTCTACCGTGTCTTTCGCCCTGGCGGAGCTTGCACAAAGGACCTTCGAGCTCGGAATCG	5796
QY	2401	TCGGCGGTGACAGGGGACGGGCAACCGGCGCTCTGACCAACCTCCGAGAGAGGGGGA	2460
Db	5797	TCGGCGGTGACAGGGGACGGGCAACGGGCTCTCTGACAGGCTTCCGAGAGAGGGGAC	5856
QY	2461	GCAAGATCTGACGTTGAGTGTATTTCTTCATGCCCCCTTTGAGGGGGAGCCGGGGAC	2520
Db	5857	GCGGAGTCTGACGTTGAGTGTATTTCTTCATGCCCCCTTTGAGGGGGAGCCGGGGAT	5916
QY	2521	CCCGATCTCAGAGCAGGGTCTTTGGTCTACCGGTAGTGAAGGCGCGGTGAGGACGTGTC	2580
Db	5917	CCCGATCTCAGAGCAGGGTCTTTGGTCTACCGGTAGTGAAGGCGCGGTGAGGACGTGTC	5976
QY	2581	TGCGTCTGAGTGTCTTACATAGGACAGGCGCTCTGATCAGCCATGGCGTGGCGGAGAA	2640
Db	5977	TGCGTCTGAGTGTCTTACATAGGACAGGCGCGCTCTGATCAGCCATGGCGTGGCGGAGAA	6036
QY	2641	AGCAAGCTGCCCATCAACGCGTTGAGCAACTCTTTGCTGCGTCAACCAACATGATGTAC	2700
Db	6037	AGCAAGCTGCCCATCAACGCGTGAAGCAACTCTTTGCTGCGTCAACCAACTTGTGTAT	6096
QY	2701	GCTACCAATCCCGGAGCGCAACCGGCGAGAGAAAGTACCTTTGACAGACTGCA	2760
Db	6097	GCTACCAATCTCGGAGCGCAACCGTGGGAGAGAAAGTACCTTTGACAGACTGCA	6156
QY	2761	ATTCCTGAGCANTACTACAGAGCGTGTCTAAGAGAGAGGAGGAGGCTTCCACGTT	2820
Db	6157	GTCTCGGAGCACACTACCGGAGCGTGTCTAAGAGAGAGGAGGAGGCTTCCACAGTT	6216
QY	2821	AAGGCTAAGCTTCTATCAGTAGAGGAACCTGCAAGCTGACGCCCCCATTTGGGCAAA	2880
Db	6217	AAGGCTAATCTTCTATCCGTGGAAGGAACCTGTAACTGACGCCCCCATTTGGGCAAG	6276
QY	2881	TCATAATTGGCTATGGGGCAAGAGCGTCCGGAACCTATCCAGCAAGGCCATTAAACAC	2940

Db	6277	TCCTAAATTTGGCTATGAGGGCAAAAGAGCTCCGAACTTATCTACAGAAAGCCGTTAAACAC	6336
Oy	2941	ATCCGCTCCGTGTGGAGGACTTGTGTGAAGACACTGAAAACCAATTTGACCAACCAATC	3000
Db	6337	ATCCGCTCCGTGTGAAGGACTTCTGTGAAGACACTGAGACACCAATTTGACACCAATC	6396
Oy	3001	ATGCGAAAAAATGAGGTTTTTCTGCTGTCCACAGAGAGAGAGAGGCCCGCAAGCTCTGC	3060
Db	6397	ATGCGAAAAAATGAGGTTTTTCTGCTGTCCACAGAGAGAGAGAGGCCCGCAAGCTCTGC	6456
Oy	3061	CTTATCCGTTGCCAGACTTGGGGGGTCCGATGTGCGAGAAAAATGGCCCTTATGACGCTG	3120
Db	6457	CTTATCCGTTGCCAGATTGGGGGGTTCGATGTGCGAGAAAAATGGCCCTTATGAGTGTG	6516
Oy	3121	GTCTCAACCCCTCCCTCAGGCTGTGATGGGCTCCTCGTACGGAATTCAGATATTCCTCTGA	3180
Db	6517	GTCTCAACCCCTCCCTCAGGCGGTGATGGGCTCTTATACGGAATTCGAATCTCTCTGGA	6576
Oy	3181	CAGCGGGTCCGAGTTTCTGTGTGAACGCTCGGAATCAAAGAGAACCCCTATGGGCTTTGCA	3240
Db	6577	CAGCGGGTCCGAGTTTCTGTGTGAACGCTCGGAAGCGAAGAAATGCCCTATGGGCTTTGCA	6636
Oy	3241	TATGACACCGCGTGTTTTGACTCAACAGTCACTGAGAAATGACATCCGTTGAGAGAGTCA	3300
Db	6637	TATGACACCGCGTGTTTTGACTCAACGCTCATGAGAAATGACATCCGTTGAGAGATCA	6696
Oy	3301	ATTATGATATGTTGTGACTTGGCCCCCGAAGCCGAGCAGGCGCATMAAGTGGCTCACAG	3360
Db	6697	ATTATGATATGTTGTGACTTGGCCCCCGAAGCCGAGCAGGCGCATMAAGTGGCTCACAG	6756
Oy	3361	CGGCTTTATATCGGGGGGTCCCTCTGACTCTAATTTCAAAGGGCGAAGCTCGGCTATCGCCG	3420
Db	6757	CGGCTTTATATCGGGGGGGCCCCCTGACTCTAATTTCTMAAGGGGAGAACTCGGCTATCGCCG	6816
Oy	3421	TGCCGCGGAGCGCGCGTGTGAAGACTAGCGTGAATACCCCTCATGTTACTTGAG	3480
Db	6817	TGCCGCGGAGCGCGTGTGAAGACTAGCGAGCAAGCTGCGGTATATACCTCATATGTTACTTGAG	6876
Oy	3481	GCCCTTGCAGCTGTGAGCTGCAAGCTGCAGAACTGCAGATGCTCGTGTGCGAGAC	3540
Db	6877	GCGGCTGCGGCGTGTGAGCTGCGAAGCTGCAGAACTGCAGATGCTCGTATGCGAGAC	6936
Oy	3541	GGGCTTGTCTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGGCGGAGCTTACAGTCTC	3600
Db	6937	GACCTTGTCTTATCTGTGAAAGGCGGGAACCCMAAGAGACAGAGGAGACTTACCGGGCC	6996
Oy	3601	TTTACGAGAGGCTATGACTAGGTACTTGACCCCCCGGGAGACCGGCCCAACAGAAATAC	3660
Db	6997	TTTACGAGAGGCTATGACTAGGTACTTGACCCCCCTTGGGAGACCGGCCCAACAGAAATAC	7056
Oy	3661	GACCTGAGTTGATATCATCATGTCTCTCCATATGTGTGCGTGTGCGACAGATGACTTGGC	3720
Db	7057	GACCTGAGTTGATATCATCATGTCTCTCCATATGTGTCAATGCGGCGACAGATGACTTGGC	7116
Oy	3721	AAAAAGGTATATCTACCTCACCCCTGAGACC	3749
Db	7117	AAAAAGGTATATCTACCTCACCCCTGAGACC	7145

Search completed: February 21, 2005, 15:39:02  
Job time : 15721 secs

Job time : 15721 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 09:01:42 : Search time 1861 Seconds  
(without alignments)  
11928.549 Million cell updates/sec

Title: US-09-664-363-22

Perfect score: 3750

Sequence: 1 TGGAGGGGGCTTCAACAGG.....ACTACTCACCCTGACCCG 3750

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1808:\*  
2: geneseqn19908:\*  
3: geneseqn20008:\*  
4: geneseqn20018:\*  
5: geneseqn20028:\*  
6: geneseqn20038:\*  
7: geneseqn20048:\*  
8: geneseqn20058:\*  
9: geneseqn20068:\*  
10: geneseqn20078:\*  
11: geneseqn20088:\*  
12: geneseqn20098:\*  
13: geneseqn20108:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3746.8	99.9	3750	2	AAQ12241 Encodes p
2	3292.2	87.8	8642	6	ABK88574 Hepatitis
3	3291.4	87.8	7987	6	AAQ25321
4	3291.4	87.8	7989	6	AAK98968 Hepatitis
5	3291.4	87.8	7989	6	AAQ25322 Hepatitis
6	3291.4	87.8	7989	12	ADJ57845
7	3291.4	87.8	7992	6	AAI47276
8	3291.4	87.8	8001	3	AAK98967 Hepatitis
9	3291.4	87.8	8637	3	AAK98966 Hepatitis
10	3291.4	87.8	8638	3	ABK88573 Hepatitis
11	3291.4	87.8	8639	6	ABK88572 Hepatitis
12	3291.4	87.8	8649	3	AAK98969 Hepatitis
13	3291.4	87.8	9605	6	ABK91424 Hepatitis
14	3291.4	87.8	9605	6	ABK91411 Hepatitis
15	3291.4	87.8	9605	6	ABK91425 Hepatitis
16	3291.4	87.8	9605	6	ABK91426 Hepatitis
17	3291.4	87.8	10690	6	ABK91412 Hepatitis
18	3291.4	87.8	10690	6	ACA61697 Hepatitis
19	3291.4	87.8	10690	10	ADC83762 PHCVNeot17
20	3291.4	87.8	10691	6	ABK91423 Hepatitis

21	3291.4	87.8	11076	3	AAK98965 Hepatitis
22	3291.4	87.8	11313	12	ADP86264
23	3291.4	87.8	11313	12	ADP86271 Hepatitis
24	3289.8	87.7	7141	6	AAQ25333
25	3289.8	87.7	7789	6	AAQ25330
26	3289.8	87.7	7987	6	AAQ25329 Hepatitis
27	3289.8	87.7	7987	6	AAQ25324 Hepatitis
28	3289.8	87.7	7989	6	AAQ25326 Hepatitis
29	3289.8	87.7	7989	6	AAQ25325
30	3289.8	87.7	7989	10	ADD93734
31	3289.8	87.7	7991	6	AAI47279 Hepatitis
32	3289.8	87.7	7992	6	AAI47280 Hepatitis
33	3289.8	87.7	7992	6	AAI47277 Hepatitis
34	3289.8	87.7	7992	6	AAI47281
35	3289.8	87.7	8732	10	ADK48121
36	3289.8	87.7	9605	6	ABK91431
37	3289.8	87.7	9605	6	ABK91429 Hepatitis
38	3289.8	87.7	9605	6	ABK91432 Hepatitis
39	3289.8	87.7	9605	6	ABK91430
40	3289.8	87.7	9605	6	ABK91428 Hepatitis
41	3289.8	87.7	9605	6	ABK91433
42	3289.8	87.7	9605	6	AAQ25332
43	3289.8	87.7	10690	6	ABK91448 Hepatitis
44	3289.8	87.7	10690	6	ABK91435 Hepatitis
45	3289.8	87.7	10690	6	ABK91243 Hepatitis

## ALIGNMENTS

RESULT 1  
ID AAQ12241 standard; cDNA; 3750 BP.  
XX  
XX AAQ12241;  
AC  
XX  
XX  
DT 25-MAR-2003 (revised)  
DT 17-SEP-1991 (first entry)  
XX  
XX Encodes portion of PT-NANBH viral non-structural protein.  
DE  
XX  
XX post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.  
KM  
XX  
OS Non-A.  
OS non-B hepatitis virus.  
OS  
XX  
XX GB2239245-A.  
XX  
PD 26-JUN-1991.  
XX  
XX 17-DEC-1990; 90GB-00027250.  
XX  
XX 18-DEC-1989; 89GB-00028562.  
PR 27-FEB-1990; 90GB-00004414.  
PR 03-MAR-1990; 90GB-00004814.  
XX  
XX (WELL ) WELLCOME FOUND LTD.  
PA (HIGH/) HIGHFIELD P E.  
XX  
XX  
PI Highfield PE, Rodgers BC, Tedder RS, Barbara JAV;  
XX  
XX WPI; 1991-187584/26.  
DR P-PSDB; AARI2599.  
DR  
XX  
XX Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA  
PT and antibodies used in diagnostic assays and in vaccines.  
XX  
XX Claim 10; Page 88-97; 108pp; English.  
XX  
XX This sequence probably encodes viral non-structural proteins of the PT-  
CC NANBH viral genome which are antigenic. It was isolated from serum of  
CC humans infected by the virus. See also AAQ12236-40 and AAQ12242. (Updated  
CC on 25-MAR-2003 to correct PA field.)

XX Sequence 3750 BP; 794 A; 1140 C; 1072 G; 744 T; 0 U; 0 Other;  
SQ Query Match 99.9%; Score 3746.8; DB 2; Length 3750;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3748; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGGAGAGGCGCTCTTCAAGGCGCTCACCCAGTGGATGCCACTTCTGTCCTCCAAACAAAG 60  
DB 1 TGGAGAGGCGCTCTTCAAGGCGCTCACCCAGTGGATGCCACTTCTGTCCTCCAAACAAAG 60  
QY 61 CAGGACAGAGACAATTTCCCTTACTGTGGCGCTACAGGCTACTGTGTGCGCTAGGCGCC 120  
DB 61 CAGGACAGAGACAATTTCCCTTACTGTGGCGCTACAGGCTACTGTGTGCGCTAGGCGCC 120  
QY 121 CAGGCCCCCCTTCATCATGAGGATCAAAATGTGAAGTGTCTCATAGGCTTAAGGCTACT 180  
DB 121 CAGGCCCCCCTTCATCATGAGGATCAAAATGTGAAGTGTCTCATAGGCTTAAGGCTACT 180  
QY 181 CTGCGGCGGCAACACCTTGTCTGTATAGGCTGGAGCCGTCCAAAACGAGTCAACCTTC 240  
DB 181 CTGCGGCGGCAACACCTTGTCTGTATAGGCTGGAGCCGTCCAAAACGAGTCAACCTTC 240  
QY 241 ACAACCCCAATAACAAATTCATCATGAGCATGATGATGATGATGATGATGATGATGATG 300  
DB 241 ACAACCCCAATAACAAATTCATCATGAGCATGATGATGATGATGATGATGATGATGATG 300  
QY 301 AGCACTGGGTGTGTGGTGGGCGGGGCTCTTGAGGCTCTGGCTGCTGATGCTTGAACA 360  
DB 301 AGCACTGGGTGTGTGGTGGGCGGGGCTCTTGAGGCTCTGGCTGCTGATGCTTGAACA 360  
QY 361 GGCAGCGTGTCTATTGTGGGTAGATCATTTGTCCGCGCGCGCGCTATTGTTCCGAC 420  
DB 361 GGCAGCGTGTCTATTGTGGGTAGATCATTTGTCCGCGCGCGCGCTATTGTTCCGAC 420  
QY 421 AGGGAAGTCTCTTACCAAGAGGTTCGATGATGATGATGATGATGATGATGATGATGATG 480  
DB 421 AGGGAAGTCTCTTACCAAGAGGTTCGATGATGATGATGATGATGATGATGATGATGATG 480  
QY 481 ATGACACAGGGAATGACAGCTCGCGAGCACTTCAAGCAAAACGCTCGGCTGGTGGCAG 540  
DB 481 ATGACACAGGGAATGACAGCTCGCGAGCACTTCAAGCAAAACGCTCGGCTGGTGGCAG 540  
QY 541 ACAGCAACCAAGAGGAGGCGCGCTGCTCCGCTGGTGGATGCAAGTGGCGAGCCCTT 600  
DB 541 ACAGCAACCAAGAGGAGGCGCGCTGCTCCGCTGGTGGATGCAAGTGGCGAGCCCTT 600  
QY 601 GAGAAGTCTTGGGCGAAACCATGTGGAATCTTATCATGCGGAGTACAGTACTTAGCAGGC 660  
DB 601 GAGAAGTCTTGGGCGAAACCATGTGGAATCTTATCATGCGGAGTACAGTACTTAGCAGGC 660  
QY 661 TTGTTCATCTGCTGGGAAATCCCGGATGTCATCTGATGCGTTTCAAGGCTCTGTC 720  
DB 661 TTGTTCATCTGCTGGGAAATCCCGGATGTCATCTGATGCGTTTCAAGGCTCTGTC 720  
QY 721 ACTAGCCCGCTCAACCAATCTAATCTCTGCTTAACTCTGCGGAGGATGAGTAAAGC 780  
DB 721 ACTAGCCCGCTCAACCAATCTAATCTCTGCTTAACTCTGCGGAGGATGAGTAAAGC 780  
QY 781 GGCCTAATGCTCTCCCGAGTGTCTTCACTTGTGAGGCGCGGATGCTGTGTGCG 840  
DB 781 GGCCTAATGCTCTCCCGAGTGTCTTCACTTGTGAGGCGCGGATGCTGTGTGCG 840  
QY 841 GCTGTGTGAGCATAGGCTTTGGGAAAGTGTCTTGAACATCTTGGGCGGCTATGAGCA 900  
DB 841 GCTGTGTGAGCATAGGCTTTGGGAAAGTGTCTTGAACATCTTGGGCGGCTATGAGCA 900  
QY 901 GAGTGGCAGGCGGCTGATGAGCTTTAAGTATGAGCGGCAAAATGCTTCAACCGAG 960  
DB 901 GAGTGGCAGGCGGCTGATGAGCTTTAAGTATGAGCGGCAAAATGCTTCAACCGAG 960  
QY 961 GACCTGATTAATTAATCTTCCCTGCAATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 961 GACCTGATTAATTAATCTTCCCTGCAATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

DB 961 GACCTGATTAATTAATCTTCCCTGCAATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
QY 1021 TGGCAGCAGATTAATGCTGCTGCGACGAGTGGTCCAGGAGAGGGGCTGTGATGATGAAC 1080  
DB 1021 TGGCAGCAGATTAATGCTGCTGCGACGAGTGGTCCAGGAGAGGGGCTGTGATGATGAAC 1080  
QY 1081 CGGCTGATTAATGCTTCCCTGCGCGGGGATCAATGTTTCCCGACGCACTATGTGCCAGAG 1140  
DB 1081 CGGCTGATTAATGCTTCCCTGCGCGGGGATCAATGTTTCCCGACGCACTATGTGCCAGAG 1140  
QY 1141 AGCGAGCGCGCAGACAGTGTCTCATGATCTCTCCAGCTTAACTAATCAACCACTGTG 1200  
DB 1141 AGCGAGCGCGCAGACAGTGTCTCATGATCTCTCCAGCTTAACTAATCAACCACTGTG 1200  
QY 1201 AAGAGGCTCCACAGTGAATTAACAGAGACTGCTCAACGCGCTGCTCGGCTGATGAGCTA 1260  
DB 1201 AAGAGGCTCCACAGTGAATTAACAGAGACTGCTCAACGCGCTGCTCGGCTGATGAGCTA 1260  
QY 1261 AAGGATGTTTGGGACTGATATGCAACATTTTGGCTGACTTCAAGACTGGCTCCAGTCC 1320  
DB 1261 AAGGATGTTTGGGACTGATATGCAACATTTTGGCTGACTTCAATTAATCTGGCTCCAGTCC 1320  
QY 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGATCAAGAGG 1380  
DB 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGATCAAGAGG 1380  
QY 1381 GTCTGGCGGAGAGACGAGCATCATGACAGCACTGTCTCATGTGAGACAGATCAACCGGA 1440  
DB 1381 GTCTGGCGGAGAGACGAGCATCATGACAGCACTGTCTCATGTGAGACAGATCAACCGGA 1440  
QY 1441 CATGTCAAAAACGCTTCATGAGGATGTTGGGCTTAAGACTGTATGTAACATGTGGCAT 1500  
DB 1441 CATGTCAAAAACGCTTCATGAGGATGTTGGGCTTAAGACTGTATGTAACATGTGGCAT 1500  
QY 1501 GGAACATTTCCCATCAAGCATTAACCAAGGCGCTTGAAGCGCTCCCGACGCGCAAC 1560  
DB 1501 GGAACATTTCCCATCAAGCATTAACCAAGGCGCTTGAAGCGCTCCCGACGCGCAAC 1560  
QY 1561 TATTCAGAGCGCTGTGGGAGTGGCTGTGAGAGATGATGAGAGTATGAGAGTATGAGAGT 1620  
DB 1561 TATTCAGAGCGCTGTGGGAGTGGCTGTGAGAGATGATGAGAGTATGAGAGTATGAGAGT 1620  
QY 1621 GATTTCCACTACGTGACAGCATGACCACTGACAAAGTAAATGCGCGTGGCAAGTTTCA 1680  
DB 1621 GATTTCCACTACGTGACAGCATGACCACTGACAAAGTAAATGCGCGTGGCAAGTTTCA 1680  
QY 1681 GCCCCGGAATTTCTTCAAGAGTGGATGGGCTGCTCAAGATGCTCCGCGTCC 1740  
DB 1681 GCCCCGGAATTTCTTCAAGAGTGGATGGGCTGCTCAAGATGCTCCGCGTCC 1740  
QY 1741 AAACCTCTCTACGGGAGAGGTCATCTCAAGTGGGCTCAACCAATTAATCTGTTGGG 1800  
DB 1741 AAACCTCTCTACGGGAGAGGTCATCTCAAGTGGGCTCAACCAATTAATCTGTTGGG 1800  
QY 1801 TGCAGACTCCATGCGAGGCCGAAACCGATGTAGAGTGTGACTCATTTCAATGCTCACGAC 1860  
DB 1801 TGCAGACTCCATGCGAGGCCGAAACCGATGTAGAGTGTGACTCATTTCAATGCTCACGAC 1860  
QY 1861 CCTTCCCATATCAACAGAGAGGCTTAAGCGGAGGCTGGCCAGGAGGCTCTCCCGCTCC 1920  
DB 1861 CCTTCCCATATCAACAGAGAGGCTTAAGCGGAGGCTGGCCAGGAGGCTCTCCCGCTCC 1920  
QY 1921 TTGGCAGGCTCTTACGCTAGCCAGTGTGTGCGGCTTCTCTGAGGCGCATTAATTAAC 1980  
DB 1921 TTGGCAGGCTCTTACGCTAGCCAGTGTGTGCGGCTTCTCTGAGGCGCATTAATTAAC 1980  
QY 1981 CAAATGACTTCCACAGACGCTGACTCATGAGGCCAACCTCTGTGGCGGATGAGATG 2040  
DB 1981 CAAATGACTTCCACAGACGCTGACTCATGAGGCCAACCTCTGTGGCGGATGAGATG 2040  
QY 2041 GGCAGGAGCATTAACCGCGCTGAGTCAAGAACAAAGTATGATCTTGAATCTTTGAC 2100  
DB 2041 GGCAGGAGCATTAACCGCGCTGAGTCAAGAACAAAGTATGATCTTGAATCTTTGAC 2100



[illegible]

Db	4526	ATCGAAGCGGGAATGCAAGCTGCGGCAACATTCTTCAACAGAAAGCAATCGGGTTGCTGCA	4585
Qy	541	ACAGCCACCAAGCAGGAGGCGCGCTGCCGTGAGAGTCAAGTGCAGCCCTT	600
Db	4586	ACAGCCACCAAGCAGGAGGCGCTGCTGCCGTGAGAGTCAAGTGCAGCCCTC	4645
Qy	601	GAGACCTTCTGGGGGAAACACATGTGGAATTCTTACGCGGGATACGTACTTAAAGCAGC	660
Db	4646	GAAACCTTCTGGGGGAAACATGTGGAATTTCTACGCGGGATACCAATTTTAAAGCAGC	4705
Qy	661	TTGTCCACTCGCTCGGGAATCCCGCATTTGCATCTGATAGGCGTTCAAGCCTCTGTC	720
Db	4706	TTGTCCACTCGCTCGGGAATCCCGCATTTGCATCTGATAGGCGATTCAAGCCTCTATC	4765
Qy	721	ACTAGCCCGCTCACCAACCCCAATCTAACCTCCTGCTTAAACTCTGGGGGAGTGGGTAACC	780
Db	4766	ACGAGCCCGCTCACCAACCCCAATACCTCCTGTTTAAACCTCGGGGGAGTGGGTAAGCC	4825
Qy	781	GCCCAACTCGCTCCCCCAAGTGTGCTTCAAGCTTTCGTAGGCGCGGCATTTGCTGTGTGC	840
Db	4826	GCCCAACTCGCTCTCCCAAGGCGCTGCTTGTGCTTCTTAAAGCGCGGCATTCGCTGTAGCGC	4885
Qy	841	GCTGTGGCAGCATAGGCGCTTGGGAAGGTGTTGTGACAATCTTGGCGGGCTTATGAGCA	900
Db	4886	GCTGTGGCAGCATAGGCGCTTGGGAAGGTGTTGTGGAATTTTGGCAGGTTATGAGCA	4945
Qy	901	GGAATGGCAGGCGCGCTCGTGGCCCTTTAAAGTCAAGCAGCGCAAAATGCCCTCACCGAG	960
Db	4946	GGAATGGCAGGCGCGCTCGTGGCCCTTTAAAGTCAAGCAGCGCAAAATGCCCTCACCGAG	5005
Qy	961	GACCTGGTTAACTTACTCCCTGTCATCTCTCTGCTGAGTCCCTGCTGTGCGGGGTCTGT	1020
Db	5006	GACCTGGTTAACTTACTCCCTGTCATCTCTCTGCTGAGTCCCTGCTGTGCGGGGTCTGT	5065
Qy	1021	TGCCAGAGCACTGCGCTCGGCAAGTGGGTCAAGGGGAGGGGGCTGTGAGTGGATGAAC	1080
Db	5066	TGCCAGAGCACTGCGCTCGGCAAGTGGGTCAAGGGGAGGGGGCTGTGAGTGGATGAAC	5125
Qy	1081	CGGCTGATAGGCTTGCCTCGCGGGGATTAAACATGTTTCCCCACGCACTATGTGCCAGAG	1140
Db	5126	CGGCTGATAGGCTTGCCTCGCGGGGATTAAACACGCTCCCCACGCACTATGTGCCAGAG	5185
Qy	1141	AGCGACGCGCAGCAGGTGTCATCAGATCCTCTCCGACTTATCTACCCAACTGTG	1200
Db	5186	AGCGACGCGCAGCAGGTGTCATCAGATCCTCTCTATCTATACCACTCAGCTGCTG	5245
Qy	1201	AAGAGGCTCCACAGTGGATTAAAGAGAACCTGCTCCAGCGCTGCTCCGGCTGTGGCTA	1260
Db	5246	AAGAGGCTCCACAGTGGATTAAAGAGAACCTGCTCCAGCGCTGCTCCGGCTGTGGCTA	5305
Qy	1261	AGGAGTGTGGGAGCTGGATATGACAGTGTGGCTACTTCAAGACTTGAGCTCAAGTCC	1320
Db	5306	AGAGATGTTTGGGATTTGGATATGACAGGTGTGACTATTTCAAGACTTGAGCTCAAGTCC	5365
Qy	1321	AAGCTCTCGCCGCAATTACCGGAGTCCCTTTTTCATGCAACCTGTGGTAAAGAGG	1380
Db	5366	AAGCTCTCGCCGCAATTGCGGGAGTCCCTCTTTCATGTCAAAGTGGGTAAAGAGGA	5425
Qy	1381	GTCTGGCGGGGAGACGGCATATGACAGACACTGTCTCATGTGGAGACAGATCACCGGA	1440
Db	5426	GTCTGGCGGGGAGACGGCATATGCAAAACACCTGCCCATGTGGAGACAGATCACCGGA	5485
Qy	1441	CATGTCAAAAACGTTCCATAGAGATGCTTGGGCTTAAAGCTGTATTAACATGTGGCAT	1500
Db	5486	CATGTCAAAAACGTTCCATAGAGATGTTGGGCTTAAAGCTGTATTAACATGTGGCAT	5545
Qy	1501	GGAACATTTCCCATCAACGCAATACCAACGCGGCCCTGTGACGCGCTCCCAAGCGCCAAAC	1560
Db	5546	GGAACATTTCCCATTAAGGCTTAACCAACGCGGCCCTGTGACGCGCTCCCGGCGCCAAAT	5605
Qy	1561	TATTTCAAGGCGCTGTGGCGGGGTGCTGTGAGAGTACGTGAAGTTTACGCGGGTGGGG	1620

Db 5606 TATTCTAGGGGCGCTGTGGCGGGTGGCTGTGAGAGATGACGTGAGTTACCGCGGGTGGG 5665  
Qy 1621 GATTTCACATACGAGAGATGACCACTGACCAACGTAAATATCCCGTCCAGATTCCA 1680  
Db 5666 GATTTCACATACGAGAGATGACCACTGACCAACGTAAATATCCCGTCCAGATTCCA 5725  
Qy 1681 GCGCCGGAATTTCTTCAACAGAAATGAGATGGGGTGGCGTGCACAGTTACGCTCCGGCTGC 1740  
Db 5726 GCGCCGGAATTTCTTCAACAGAAATGAGATGGGGTGGCGTGCACAGTTACGCTCCAGCTGC 5785  
Qy 1741 AAACCTCTCTTCAAGGAGAGATGACATTTCCAGGTGGGGCTCAACCAATACCTGGTGGG 1800  
Db 5786 AAACCTCTCTTCAAGGAGAGATGACATTTCCAGGTGGGGCTCAACCAATACCTGGTGGG 5845  
Qy 1801 TCGAGCTCCATGCGAGCCCGAACCAGATGTAGCAATGCTCAATTCATGCTCACCGAC 1860  
Db 5846 TCAAGCTCCATGCGAGCCCGAACCAGATGTAGCAATGCTCAATTCATGCTCACCGAC 5905  
Qy 1861 CCTCCACATCAACAGAGAGACGGCTAAGCGCAGCTGGCCAGGGGGTCTCCCGCTCC 1920  
Db 5906 CCTCCACATCAACAGGGAGACGGCTAAGCGTAAAGCTGGCCAGGGGATCTCCCGCTCC 5965  
Qy 1921 TTGGCAGCTCTTCACTAGTACCACTTGTCTGCGCTTCTCGAAGGCGACATACATTACC 1980  
Db 5966 TTGGCAGCTCTTCACTAGTACCACTTGTCTGCGCTTCTCGAAGGCGACATACATTACC 6025  
Qy 1981 CAAATATGACTTCCCAAGACGCTGACCTATGAGAGCCAACTCTCTGTGGCGGATGAGATG 2040  
Db 6026 CGTATGACTTCCCAAGACGCTGACCTATGAGAGCCAACTCTCTGTGGCGGAGATG 6085  
Qy 2041 GGGCGGAGACATTAACCGCGTGAAGTCAAGAACAAAGTAGTAATCTGGAATCTTTTGAC 2100  
Db 6086 GGGCGGAGACATTAACCGCGTGAAGTCAAGAACAAAGTAGTAATCTTTTGAC 6145  
Qy 2101 CGGCTCCGAGCGGAGAGATGAGCGGGAATGTCCCTCCCGCGAGATCTTCGCGAAA 2160  
Db 6146 CGGCTCCGAGCGGAGAGATGAGCGGGAATGTCCCTCCCGCGAGATCTTCGCGAAG 6205  
Qy 2161 TCCAAAGAAATTTCCCAACAGAGATGCGCGCATGGGCAAGCCCGATTACACCTCTCGCTG 2220  
Db 6206 TCCAAAGAAATTTCCCTCGAGAGATGCCATATGGGCAAGCCCGATTACACCTCTCGCTG 6265  
Qy 2221 CTGAGAGCTTGAAGGCGCGGACTAGCTCCCTCCAGATGGTAACATGGGGTCCCACTGCA 2280  
Db 6266 TTRGAGTCTTGAAGGCGCGGACTAGCTCCCTCCAGATGGTAACATGGGGTCCCACTGCGC 6325  
Qy 2281 CCTACTAAGACCCCTCTTATACCACTTCAACGAGAGAAAGACAATTTGTTCTGACAAA 2340  
Db 6326 CTGCGCAAGGCGCCCTCGATACCACTTCAACGAGAGAAAGACAATTTGTTCTGTCAGAA 6385  
Qy 2341 TCCACCGTGTCTTGTGCGCTGGGAGCTTGGCCAAAGGCTTTCCGTAAGCTCCGAAACCG 2400  
Db 6386 TCTACCGTGTCTTGTGCGCTGGGAGCTTGGCCAAAGGCTTTCCGTAAGCTCCGTAATCG 6445  
Qy 2401 TCGGCGGTGACAGCGGCAAGCAACCGCCCTCTGACCAACCTCTCGAGCGAGGGA 2460  
Db 6446 TCGGCGGTGACAGCGGCAAGCAACCGCCCTCTGACCAACCTCTCGAGCGAGGGA 6505  
Qy 2461 GCAGATCTGACGCTTGAAGTGTATCTCTCCATGCCCCCTTGAAGGGGAGCGCGGGGAT 2520  
Db 6506 GCGGATCTGACGCTTGAAGTGTATCTCTCCATGCCCCCTTGAAGGGGAGCGCGGGGAT 6565  
Qy 2521 CCGGATCTGACGCGGCTTTGGTCTACCTGTAAGTGAAGGCGCGGTGAAGAGCTGTC 2580  
Db 6566 CCGGATCTGACGCGGCTTTGGTCTACCTGTAAGTGAAGGCGCTGATGAGAGCTGTC 6625  
Qy 2581 TGTGCTGATGATCTTCAACATGAGAGGCGCTGTATCAAGCGCATGCGTGGGAGGAA 2640  
Db 6626 TGTGCTGATGATCTTCAACATGAGAGGCGCGCTGTATCAAGCGCATGCGTGGGAGGAA 6685  
Qy 2641 AGCAAGCTGCGCATCAACGCGTGAAGCACTTTTGTGCTGCTACCAACATGCTTAC 2700  
Db 6686 ACCAAGCTGCGCATCAACGCGTGAAGCACTTTTGTGCTGCTACCAACATGCTTAT 6745

Qy 2701 GCTACCAATCCCGACGCGAGCGGACGCGAGAGAGTCACTTTGACAGACTGCA 2760  
Db 6746 GCTACCAATCTTGGACGCGAGAGCTTGGCGAGAGAGTCACTTTGACAGACTGCA 6805  
Qy 2761 ATCTGAGACATCACTACAGAGAGTGTCTCAAGAGATGAAGGCGAGGCGTCCACAGTT 2820  
Db 6806 GTCTGAGACATCACTACGCGAGAGTGTCTCAAGAGATGAAGGCGAGGCGTCCACAGTT 6865  
Qy 2821 AAGCTTGAAGTTTATCATGATGAGAGAGCTGCAAGCTGACGCGCCCACTTTGGCTTAA 2880  
Db 6866 AAGCTTGAAGTTTATCATGATGAGAGAGCTGCAAGCTGACGCGCCCACTTTGGCTTAA 6925  
Qy 2881 TCTAAATTTGGCATGAGGAGCAAGAGAGTCCGGAACCTATCCAGAGGCAATTAACAC 2940  
Db 6926 TCTAAATTTGGCATGAGGAGCAAGAGAGTCCGGAACCTATCCAGAGGCAATTAACAC 6985  
Qy 2941 ATCCGCTCCGTGTGAGAGAGCTTGTGAGAGACTGAACCAATTGACACCACTATC 3000  
Db 6986 ATCCGCTCCGTGTGAGAGAGCTTGTGAGAGACTGAACCAATTGACACCACTATC 7045  
Qy 3001 ATGCAAAAAATGAGGTTTTCTGCTGCAACAGAGAGAGGCGGCAACCACTGCTGC 3060  
Db 7046 ATGCAAAAAATGAGGTTTTCTGCTGCAACAGAGAGAGGCGGCAACCACTGCTGC 7105  
Qy 3061 CTATGCTGTCCAGACTTGGGGGCTCCGTGTGAGAGAAATGGGCTCTATGACGTG 3120  
Db 7106 CTATGCTGTCCAGACTTGGGGGCTCCGTGTGAGAGAAATGGGCTCTATGACGTG 7165  
Qy 3121 GTCTCAACCTCCCTGAGGCTGTGATGAGGCTCTCTGTACAGATTCAGTATCTCTGCA 3180  
Db 7166 GTCTCAACCTCCCTGAGGCTGTGATGAGGCTCTCTGTACAGATTCAGTATCTCTGCA 7225  
Qy 3181 CAGCGGCTGAGTTCTGTGTGAACGCTGTGAATCAAGAGAGCCCTATGGCTTTGCA 3240  
Db 7226 CAGCGGCTGAGTTCTGTGTGAATGCTGTGAAGGCAAGAAAGCCCTATGGCTTTGCA 7285  
Qy 3241 TATGACACCGCGTTTGTGATCAACAGTCACTGAAGAAATGCACTCCGTGAGAGGTCA 3300  
Db 7286 TATGACACCGCGTTTGTGATCAACAGTCACTGAAGAAATGCACTCCGTGAGAGGTCA 7345  
Qy 3301 ATTATCAATGTTGTGATCTTGGCGCCCGAAGCGAGACAGGCTAAGTGTGCTCAAGAG 3360  
Db 7346 ATTATCAATGTTGTGATCTTGGCGCCCGAAGCGAGACAGGCTAAGTGTGCTCAAGAG 7405  
Qy 3361 CGGCTTTATATCGGGGCTCCCTGACTAATTTCAAAAGGCGAGAACTGGGCTATCGCGG 3420  
Db 7406 CGGCTTTATATCGGGGCTCCCTGACTAATTTCAAAAGGCGAGAACTGGGCTATCGCGG 7465  
Qy 3421 TGGCGGCGAGCGGCTGTGAGAGACTAGTGGGTTAATACCTTCAATGTTACTTTGAAG 3480  
Db 7466 TGGCGGCGAGCGGCTGTGAGAGACTAGTGGGTTAATACCTTCAATGTTACTTTGAAG 7525  
Qy 3481 GCCTGTGACCGCTGTGAGGCTGCAAGCTCAAGAGCTGACAGATGCTGCTGTGCGAGAG 3540  
Db 7526 GCCTGTGACCGCTGTGAGGCTGCAAGCTGCAAGAGCTGACAGATGCTGCTGTGCGAGAG 7585  
Qy 3541 GGCCTTGTGCTATCTGTGAGAGCGGAGAACCCAGAGAGCGGCGAGGCTTACAGAGTC 3600  
Db 7586 GACCTTGTGCTATCTGTGAGAGCGGAGAACCCAGAGAGCGGCGAGGCTTACAGAGTC 7645  
Qy 3601 TTTACGAGAGCTTATGACTAGTACTGTGCCCCCGGGGAGCCCGCCCAACCAAGATAC 3660  
Db 7646 TTTACGAGAGCTTATGACTAGTACTGTGCCCCCGGGGAGCCCGCCCAACCAAGATAC 7705  
Qy 3661 GACCTGAGTGTATATCATATGCTCTCTCAATGTTGCTGCGGACAGATGATGCTGAC 3720  
Db 7706 GACCTGAGTGTATATCATATGCTCTCTCAATGTTGCTGCGGACAGATGATGCTGAC 7765  
Qy 3721 AAAAGGTAATCACTCAACCGGTGACC 3749  
Db 7766 AAAAGGTAATCACTCAACCGGTGACC 7794

RESULT 3  
AAD25321  
ID AAD25321 standard; cDNA; 7987 BP.  
XX  
AC AAD25321;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Hepatitis C virus (HCV) replBartMan/delta2U's cDNA.  
XX  
KM Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
KM gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
KM ss.  
XX  
OS Hepatitis C virus.  
XX  
PN MO200189364-A2.  
XX  
PD 29-NOV-2001.  
XX  
PF 23-MAY-2001; 2001MO-US016822.  
XX  
PR 23-MAY-2000; 2000US-00576989.  
XX  
PA (UNIW ) UNIV WASHINGTON.  
XX  
PI Rice CM, Blight KJ;  
XX  
DR WPI; 2002-066755/09.  
XX  
PT Hepatitis C virus variants having greater transfection efficiency and  
PT ability to survive subpassage, useful as a vaccine for immunizing primate  
PT to the virus, comprise non-naturally occurring viral sequences.  
XX  
PS Claim 44; Page 66-69; 174pp; English.  
XX  
CC The invention relates to Hepatitis C virus (HCV) variants which include  
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
CC variants that have a transfection efficiency and ability to survive  
CC subpassage greater than HCV that have wild-type polyprotein coding  
CC regions. The polynucleotides of the invention are useful for identifying  
CC a cell line that is permissive for infection with HCV and detecting  
CC replication of HCV in cells of the cell line. They are also useful for  
CC testing a compound for anti-viral properties and for inhibiting HCV  
CC infection. They are also useful for the generation of defined HCV virus  
CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
CC attachment, penetration and entry, structure/function studies on HCV  
CC proteins and RNA elements and identification of new antiviral targets, a  
CC systematic survey of cell culture systems and conditions to identify  
CC those that support wild-type and variant HCV RNA replication and particle  
CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of attenuated or  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilisation of the HCV  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
CC The present sequence is Hepatitis C virus (HCV) replBartMan/delta2U's  
CC cDNA  
XX  
SQ Sequence 7987 BP; 1647 A; 2368 C; 2243 G; 1729 T; 0 U; 0 Other;  
XX  
XX  
Query Match 87.8%; Score 3291.4; DB 6; Length 7987;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;  
QY 1 TGGAGAGGCGCTTTCACAGGCTCCACCGTGTGATGCCCACTTCTGTCCCAACAAAG 60

DB 3397 TGGAGAGGCGCTTTCACAGGCTCCACCGTGTGATGCCCACTTCTGTCCCAACAAAG 3456  
QY 61 CAGGACGAGAGACAACTTCCCTAACCTGATGAGGCTACAGGCTTACTGTGCGTAGAGGC 120  
DB 3457 CAGGACGAGAGACAACTTCCCTAACCTGATGAGGCTACAGGCTTACTGTGCGTAGAGGC 3516  
QY 121 CAGGACGAGAGACAACTTCCCTAACCTGATGAGGCTTACTGTGCGTAGAGGC 180  
DB 3517 CAGGCTCCACCTTCATCTGTGAGGCAAAATGTGGAATGTCTCATACGGCTAAAGCCTACG 3576  
QY 181 CTGCGGGGGCCAAACACCTTGTGCTGTATAGCTTGGAGCCGTCCAAACGAGGTCAACCTC 240  
DB 3577 CTGCAAGGGCCAAACGCTTGTGTATAGCTTGGAGCCGTTCAAACGAGGTCTACTAC 3636  
QY 241 ACACACCCCATTAACAAATTCATCATGAGCATGATGATGACGACCTGAGAGGTCTACG 300  
DB 3637 ACACACCCCATTAACAAATTCATCATGAGCATGATGATGACGACCTGAGAGGTCTACG 3696  
QY 301 AGCAGCTGGGTGTGTGTGGGCGGGGATCCTTGACAGCTCTGGCTGCTATTTGCTTGACACA 360  
DB 3697 AGCAGCTGGGTGTGTGTGGGCGGGGATCCTTGACAGCTCTGGCTGCTATTTGCTTGACACA 3756  
QY 361 GGCAGCGTGTGCTATTTGTGGGTAGATCATCTTGTCCGGCGCGCGGCTATTTGTTCCCGAC 420  
DB 3757 GGCAGCGTGTGCTATTTGTGGGTAGATCATCTTGTCCGGAAAGCGCGCATCATTTCCCGAC 3816  
QY 421 AGGAGAGTCTCTACACGAGAGGTTCATGATGATGATGAGAGGTGCGCTGACCTCCCTTAC 480  
DB 3817 AGGAGAGTCTCTTACCGGAGATTCATGATGATGAGAGGTGCGCTCACAACCTCCCTTAC 3876  
QY 481 ATGAGACAGGAGATGAGAGTTCGCGGACAGTTCAGCAAAAGAGCGCTCGGTTGTCGAC 540  
DB 3877 ATGAGACAGGAGATGAGAGTTCGCGGACAGTTCAGCAAAAGAGCATTCGGGTTGTCGAC 3936  
QY 541 ACAGCCACCAAGAACGAGAGCGGCTGCTCCGCTGTGTGAGATGCCAAGTGGCCAGCCCTT 600  
DB 3937 ACAGCCACCAAGAACGAGAGCGGCTGCTCCGCTGTGTGAGATGCCAAGTGGCCAGCCCTC 3996  
QY 601 GAGACCTTCTGGGCGAAACACATGTGGAACCTTATCATGAGCGGATACAGTATTAGCAGGC 660  
DB 3997 GAGACCTTCTGGGCGAAACATGTGGAACCTTATCATGAGCGGATACAGTATTAGCAGGC 4056  
QY 661 TTGTCACTCTGTGCTGGGAATCCCGGATGTGATCACTGATGAGCGTTCAAGCCTCTGTC 720  
DB 4057 TTGTCACTCTGTGCTGGGAATCCCGGATGTGATCACTGATGAGCGTTCAAGCCTCTGTC 4116  
QY 721 ACTAGCCCGCTCAACCAATATACCTTCCTGCTTAAATCTTGGGGGGAATGGGTATGCC 780  
DB 4117 ACCAGGCCGCTCAACCAATATACCTTCCTGCTTAAATCTTGGGGGGAATGGGTATGCC 4176  
QY 781 GCCCACTGCTCCCGGCAAGTGTGCTTACAGCTTGTGTAAGGCGCGCATTTGCTGTGCG 840  
DB 4177 GCCCACTGCTCCCGGCAAGTGTGCTTACAGCTTGTGTAAGGCGCGCATTTGCTGTGCG 4236  
QY 841 GCTGTGTGAGCATATAGGCTTGGGAAGTGTCTTGTGGAATCTTGTGCGGCTATGAGACA 900  
DB 4237 GCTGTGTGAGCATATAGGCTTGGGAAGTGTCTTGTGGAATCTTGTGCGGCTATGAGACA 4296  
QY 901 GGAATGAGAGCGCGCTGCTGTGAGCTTAAAGTATGAGGCGCAAAATGCCCTTACACGAG 960  
DB 4297 GGGGTGAGAGCGCGCTGCTGTGAGCTTAAAGTATGAGGCGCAAAATGCCCTTACACGAG 4356  
QY 961 GACCTGTTAACTTACTCCCTGTCATCTCTCTGCTGTGAGCGCTGAGCGTGGGTCTGTG 1020  
DB 4357 GACCTGTTAACTTACTCCCTGTCATCTCTCTGCTGTGAGCGCTGAGCGTGGGTCTGTG 4416  
QY 1021 TGGCAGAGCATATCTGCTGCGCACGTGATCAAGGAGAGGAGGCTGTGACGTGATGATAC 1080  
DB 4417 TGGCAGAGCATATCTGCTGCGCACGTGATCAAGGAGAGGAGGCTGTGACGTGATGATAC 4476  
QY 1081 CGGCTGATAGCGTTGCGCTGTGCGGGGTAAACATGTTTCCCGCAGCATATGTGCCAGAG 1140



Db 4477 CGGCTATAGCGTTTCGTCGCGGGGTAACACGCTCTCCCCACGCACTATGTGCTGAG 4536  
Qy 1141 AGGAGCGCGCAGCAGTGTCACTCAAGATCTCTCGACCTTACTATCAACCACTGTG 1200  
Db 4537 AGGAGCGCTGACAGTGTCACTCAAGATCTCTCAAGTCTTACCACTCACTGCTG 4596  
Qy 1201 AAGAGCTCCACGAGTGAATTAACGAGACTGCTCCACGCTCTCCGCTGCTGCTA 1260  
Db 4597 AAGAGCTTCAACAGTGAATTAACGAGACTGCTCCACGCTCTCCGCTGCTGCTA 4656  
Qy 1261 AAGGATGTTGGAGATGATATGCAAGTGTGGCTGACTTAAGACTGCTGCTGCTGCT 1320  
Db 4657 AGAGATGTTGGAGATGATATGCAAGTGTGGCTGACTTAAGACTGCTGCTGCTGCT 4716  
Qy 1321 AAGCTCTGCGGAGTATTAACCGGAGTCCCTTTTCTCATGCTCAAGCTGGGTCAAGAGG 1380  
Db 4717 AAGCTCTGCGGAGTATTAACCGGAGTCCCTTTTCTCATGCTCAAGCTGGGTCAAGAGG 4776  
Qy 1381 GTCTGCGGGGAGAGCGGATCATAGCAGACCACTGCTCATGTGAGCAGATCACCGA 1440  
Db 4777 GTCTGCGGGGAGAGCGGATCATAGCAGACCACTGCTCATGTGAGCAGATCACCGA 4836  
Qy 1441 CATGTCAAAAACGTTTCATAGAGATGTTGGCTTAAAGCTGTAGTAACTATGTCAT 1500  
Db 4837 CATGTCAAAAACGTTTCATAGAGATGTTGGCTTAAAGCTGTAGTAACTATGTCAT 4896  
Qy 1501 GGAACATTTCCCATCAACCGATACACACGCGGCTCCGTCAGCGCTCCCGAGCGCAAC 1560  
Db 4897 GGAACATTTCCCATCAACCGATACACACGCGGCTCCGTCAGCGCTCCCGAGCGCAAC 4956  
Qy 1561 TATTCAGGGCGCTGTGCGGGTGTGCTGTGAGAGTACGTGAGGTTACGCGGGTGGG 1620  
Db 4957 TATTCAGGGCGCTGTGCGGGTGTGCTGTGAGAGTACGTGAGGTTACGCGGGTGGG 5016  
Qy 1621 GATTTCACTACGTGACGAGCATGACCACTGACACGTAATAATGCCGTGACAGTTCCA 1680  
Db 5017 GATTTCACTACGTGACGAGCATGACCACTGACACGTAATAATGCCGTGACAGTTCCA 5076  
Qy 1681 GCGCCCGAATCTTTCACAGAGTGAATGGGTGCGGTGTCACAGTACGCTTCACGCGTGC 1740  
Db 5077 GCGCCCGAATCTTTCACAGAGTGAATGGGTGCGGTGTCACAGTACGCTTCACGCGTGC 5136  
Qy 1741 AAACCTCTCCTACGGGAGAGTCACTTCAGAGTGGGCTCAACCAATACTGTTGGG 1800  
Db 5137 AAACCTCTCCTACGGGAGAGTCACTTCAGAGTGGGCTCAACCAATACTGTTGGG 5196  
Qy 1801 TCGCAGCTCCCATGCGAGCGGACCGAGTGAAGTGTCACTTCATGCTCACCGAC 1860  
Db 5197 TCGCAGCTCCCATGCGAGCGGACCGAGTGAAGTGTCACTTCATGCTCACCGAC 5256  
Qy 1861 CCTTCCCATCAACAGAGAGAGGCTTAAGCGAGCTGGGCGAGGGGTCTCCCTCTCC 1920  
Db 5257 CCTTCCCATCAACAGAGAGAGGCTTAAGCGAGCTGGGCGAGGGGTCTCCCTCTCC 5316  
Qy 1921 TTGGCAGGCTTTCAGTAGCAGTGTGTCGCGCTTCTCGAAGGCGACATACATTAC 1980  
Db 5317 TTGGCAGGCTTTCAGTAGCAGTGTGTCGCGCTTCTCGAAGGCGACATACATTAC 5376  
Qy 1981 CAATAATGACTTCCAGACGCTGACCTATGAGCGCAACCTCTGTGGCGCATGAGATG 2040  
Db 5377 CGTATGACTTCCCGAGCGCTGACCTATGAGCGCAACCTCTGTGGCGCGAGAGATG 5436  
Qy 2041 GGGGGGAGCATTTACCGCGGTGAGTCAAGAGACAAAGTATATCTGAGCTTTTGAC 2100  
Db 5437 GGGGGGAGCATTTACCGCGGTGAGTCAAGAGACAAAGTATATCTGAGCTTTTGAC 5496  
Qy 2101 CGGCTCCAGCGGAGAGTGAAGCGGAAAGTGTCCGCTCCGGGAGGATCTGCGGAAA 2160  
Db 5497 CGGCTCCAGCGGAGAGTGAAGCGGAAAGTGTCCGCTCCGGGAGGATCTGCGGAAA 5556  
Qy 2161 TCCAGAAATTTCCACACGAGATGCGGCAATGGGCAAGCGCGGATTAACAACCTCGCTG 2220  
Db 5557 TCCAGAAATTTCCCTGAGGAGATGCCAATATGGGCAAGCGCGGATTAACAACCTCGCTG 5616

Qy 2221 CTGAGTCTCGAAGGCGCCGAGTACTGCTCCCTCAAGTGTACATGGTGCCTCACTGCA 2280  
Db 5617 TTAGATCTCTGAGAGAGACCGGACATGCTCCCTCAAGTGTACATGGTGCCTCACTGCA 5676  
Qy 2281 CCTACTAAGACCTCTCTATATACCACTCAACGAGAGAGAGACAGTGTGTTGACAGAA 2340  
Db 5677 CTTGCAAGGCGCTCCGATTAACACTTCAACGAGAGAGAGACAGTGTGTTGTTGACAGAA 5736  
Qy 2341 TCCACGCTGCTTTCCTGCGGAGGCTTGGCCAGAAAGGCTTGGTGTAGCTCCGAAACG 2400  
Db 5737 TCTACGCTGCTTTCCTGCGGAGGCTTGGCCAGAAAGGCTTGGTGTAGCTCCGAAACG 5796  
Qy 2401 TCGGCGGTGACAGCGGACCGGCAACGCGCTCTGACCAACCTTCCGACGAGCGGAG 2460  
Db 5797 TCGGCGGTGACAGCGGACCGGCAACGCGCTCTGACCAACCTTCCGACGAGCGGAG 5856  
Qy 2461 GCAGATCTGACGTTGAGTGTATTTCTCATGCCCCCTTGAAGGGAGCGCGGGAGC 2520  
Db 5857 GCGGATCCGAGGTTGATGCTCTCTCATGCCCCCTTGAAGGGAGCGCGGGAGC 5916  
Qy 2521 CCGGATCTCAGGAGCGGCTTGTGTCTACCGTGAAGAGAGCGGTGAGACGTCTGTC 2580  
Db 5917 CCGGATCTCAGGAGCGGCTTGTGTCTACCGTGAAGAGAGCGGTGAGACGTCTGTC 5976  
Qy 2581 TGCTGCTCGATGCTCTACACATGAGCAGGCGCTGTATGACGCGCATGCGTGGAGAGAA 2640  
Db 5977 TGCTGCTCGATGCTCTACACATGAGCAGGCGCTGTATGACGCGCATGCGTGGAGAGAA 6036  
Qy 2641 AGCAGCTGCCATCAACCGGTTGAGCAACTTTTGTGCTGCTGACCAACAATGCTTAC 2700  
Db 6037 ACCAAGCTGCCATCAACCGGTTGAGCAACTTTTGTGCTGCTGACCAACAATGCTTAC 6096  
Qy 2701 GCTACCAATCCCGACGCGCAAGCGGCGAGAGAGTCACTTTGACAGCTGCA 2760  
Db 6097 GCTACCAATCCCGACGCGCAAGCGGCGAGAGAGTCACTTTGACAGCTGCA 6156  
Qy 2761 ATCCGGAAGATGCTACACAGGACGTCCTCAAGAGATGAAGGCGAAGGCTGCAAGTT 2820  
Db 6157 GTCTGGAAGACCACTACGAGGAGTGTCTGAGAGATGAAGGCGAAGGCTGCAAGTT 6216  
Qy 2821 AAGGCTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACGCCCCCACTTGGGCAAA 2880  
Db 6217 AAGGCTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACGCCCCCACTTGGGCAAA 6276  
Qy 2881 TCTAAATTTGGCTATGCGGCAAGAGCTGCCGAACCTTATCCAGCAAGGCTTAACTAC 2940  
Db 6277 TCTAAATTTGGCTATGCGGCAAGAGCTGCCGAACCTTATCCAGCAAGGCTTAACTAC 6336  
Qy 2941 ATCCGCTCGGTGGGAGGACTTGTGGAAGCACTGAACCAATTGACACCACTATC 3000  
Db 6337 ATCCGCTCGGTGGGAGGACTTGTGGAAGCACTGAACCAATTGACACCACTATC 6396  
Qy 3001 ATGGCAAAAATGAGGTTTCTGCGTCAACAGAGAGAGAGCGGCAAGCGAGCTGCG 3060  
Db 6397 ATGGCAAAAATGAGGTTTCTGCGTCAACAGAGAGAGAGCGGCAAGCGAGCTGCG 6456  
Qy 3061 CTTATGCTTCCAGACTTGGGGGTGCGGTGTGCGAGAGAAATGGCCCTTATGACGTC 3120  
Db 6457 CTTATGCTTCCAGACTTGGGGGTGCGGTGTGCGAGAGAAATGGCCCTTATGACGTC 6516  
Qy 3121 GTCTCACCTCCCTGAGGCTGATGAGGCTCCTGTAGAGATTCAGATTTCTCTGGA 3180  
Db 6517 GTCTCACCTCCCTGAGGCTGATGAGGCTCCTGTAGAGATTCAGATTTCTCTGGA 6576  
Qy 3181 CAGCGGCTGAGTTCCTGTGTAACGCTGGAATCAAGAGAGACCCCTATGAGCTTTGCA 3240  
Db 6577 CAGCGGCTGAGTTCCTGTGTAATGCTGGAATGCTGGAAGAGAGAAATGCTATGAGCTT 6636  
Qy 3241 TATGACACCGGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTATGAGAGTCA 3300  
Db 6637 TATGACACCGGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTATGAGAGTCA 6696

OY	3301	TTTATTCATGTTGACTTTSACCCCGAGACGACGAGCCATTAAGTGCCTCAGAG	3360
Db	6657	ATTCACCATGTTGACTTTSACCCCGAGAGCCAGACGCATTAAGTGCCTCAGAG	6756
OY	3361	CGGCTTATATCGAGGGTCCCTCTGACTAATTCAAAAGGCGAGAACTGCGGCTATCGCGG	3420
Db	6757	CGGCTTATATCGAGGGGCCCCCTGCTAATTCTAAAGGCGAGAACTGCGGCTATCGCGG	6816
OY	3421	TGCGCGGAGACGGGCTGCTGACATAGTACGTGCGGTATACCCATCACTGTACTTAAG	3480
Db	6817	TGCGCGGAGAGGGTACTGACACGACGCTGCGGTATATCCCTCACTGTACTTAAG	6876
OY	3481	GCTCTGCAGCCTGTGCAGCTGCAGAAAGCTCCAGACTGCAGATGCTTGTGTCGAGAC	3540
Db	6877	GCGCTGCGGCTGTGCAGCTGCCAAGCTCCAGAGCTCAAGATGCTGTATGGGAGAC	6936
OY	3541	GCGCTTGTCTTATTTGTGAGAGCGCGGGAAACCAAGAGAGACGGCGGAGCTTACGAGTC	3600
Db	6937	GACCTTGTCTTATTTGTGAAAGCGCGGGAGCCMAAGAGAGACGAGCCAGCTTACGGGCC	6996
OY	3601	TTCAACGAGGCGTATGATCAGGTAATCTGTCCCCCGGAGACCGGCCCAACCAAGATAC	3660
Db	6997	TTCAACGAGGCGTATGATCAGATATCTTCTCCCCCTGTGGAGACCGGCCCAACCAAGATAC	7056
OY	3661	GACCTGAGGTTGATTAACATCAATGCTCTTCATATGTTGAGTCGGCGACGATGACTTGGC	3720
Db	7057	GACCTGAGGTTGATTAACATCAATGCTCTTCATATGTTGAGTCGGCGACGATGACTTGGC	7116
OY	3721	AAAAGGTAATACCTCAACCCGTCAGCC	3749
Db	7117	AAAAGGTAATACCTCAACCCGTCAGCC	7145

RESULT 4  
 AAA98968  
 ID AAA98968 standard; DNA; 7989 BP.  
 XX  
 AC AAA98968;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Hepatitis C virus DNA fragment SEQ ID NO: 4.  
 XX  
 KW Cell culture; therapy; infection; vaccine; diagnosis; gene therapy; ds.  
 XX  
 CS Hepatitis C virus.  
 XX  
 PN DE19915178-A1.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 03-APR-1999; 99DE-01015178.  
 XX  
 PR 03-APR-1999; 99DE-01015178.  
 XX  
 PA (UWMA-) UNIV MAINZ GUTENBERG JOHANNES.  
 XX  
 PI Bartenschlager R;  
 XX  
 DR WPI; 2000-629140/61.  
 XX  
 PT Cell culture system for hepatitis C virus, useful e.g. in screening for  
 PT therapeutic agents, comprises human hepatoma cells containing a viral RNA  
 PT construct that includes a selectable gene.  
 XX  
 PS Claim 8; Page 37-43; 58pp; German.  
 CC  
 CC This invention describes a novel Hepatitis C virus (HCV) cell culture  
 CC system comprising human hepatoma cells that contain an integrated HCV-RNA  
 CC construct (I). (I) contains the HCV-specific RNA segments 5'-NTR (non-  
 CC translated region), NS (non-structural)3, NS4A, NS4B, NS5A, NS5B and 3'-  
 CC NTR, and a selectable (marker) gene (II). The cell cultures, and/or (I),  
 CC are used to prepare, evaluate and/or test therapeutic and/or diagnostic

CC agents for HCV infections, and to prepare vaccines against HCV infection  
CC (particularly preparation of attenuated HCV). The can also be used for  
CC preparation of a liver-specific delivery system for gene therapy, and to  
CC identify cells permissive for HCV replication. Virus RNA replicates  
CC autonomously and with high efficiency in this cellular system, so that  
CC variations in replication rates can be measured (for screening antiviral  
CC agents) quantitatively or qualitatively, using standard laboratory  
CC equipment. Efficient replication of HCV RNA is only achieved when the  
CC specified RNA segments are present and when the transfected cells are  
CC maintained under permanent selection pressure

XX Sequence 7969 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other;  
XQ

Query Match	87.8%;	Score 3291.4;	DB 3;	length 7989;
Best Local Similarity	92.4%;	Pred. No. 0;		
Matches 3463;	Conservative	0;	Mismatches 286;	Indels 0;
			Gaps	0

OY	1	TGGGAGGGCGCTTTCACAGGCGCTCACCGCAGTGGATGCCCATCTTCGTGCCAACAAG	60
Db	3397	TGGGAGGCGCTTTACAGGCGCTCACCATAGACGCCCATTTCTGTGCCAGACTAG	3456
OY	61	CAGGCAAGAACAACTTCCCTTACTGTGTGGCGTACAGGCTACTGTGTGGCGTA	120
Db	3457	CAGGCAAGAACAACTTCCCTTACTGTGTAGCAATACAGGCTACGGTGTGGCCAGGCGCT	3516
OY	121	CAGGCCCCACTTCATCATATGGGATCAATATGGAAAGTGTCAATACGGCTTAAAGCTACT	180
Db	3517	CAGGCTCCACTTCATCATGTGGGACCAATATGGAAAGTGTCAATACGGCTTAAAGCTTACG	3576
OY	181	CTGGCGGGGCGCAACACCTTGTGTATATAGGCTGGGAGCGCTCCAAAACGAGTCAACCTC	240
Db	3577	CTGACAGGGCCAAAGCCCTCTGTGTATATAGGTGGGAGCCGTTCAAAAACGAGTTACTAC	3636
OY	241	ACACACCCCATTAACCAATTCATATGGCATGCTACAGCCGACTGGAGTGTCTACG	300
Db	3637	ACACACCCCATTAACCAATTCATATGGCATGCTACAGCTGCTGACCTGGAGTGTCTACG	3696
OY	301	AGACCTGGGGGTGCTGGTGGGCGGGGTCCTTGACCTGGCTGGGTATTTGTGACAACA	360
Db	3697	AGACCTGGGGGTGCTGGTGGGCGGGGTCCTTGACCTTGGCTGGGTATTTGTGACAACA	3756
OY	361	GGCAGCGTGTGCTATTGTGGTAGGATCATCTTGTCCGGCGGGCCGGCTATTGTCCCGAC	420
Db	3757	GGCAGCGTGTGCTATTGTGGTAGGATCATCTTGTCCGGAACCGGCGCATATTCCCGAC	3816
OY	421	AGGGAAGTCTCTTACAGAGAGTTGATGAGATGGAAGAAGCGCGCTGCACCTCCCTTAC	480
Db	3817	AGGGAAGTCTCTTACCGGAGTTGATGAGATGGAAGAAGCGCGCTGCACCTCCCTTAC	3876
OY	481	ATTGAGCAGGGAAATGACAGCTGCGCGAGCAGTTCAAGCAAAAAGCGCTCGGGTTCTCGAG	540
Db	3877	ATTGAGCAGGGAAATGACAGCTGCGCGAACATTTCAACAGAAAGCAATCGGGTTCTCGAA	3936
OY	541	ACAGCCACCAAGCAAGCGGAGGCGCTGCTCCGTTGGTAGTCCAAATGGCGAGCCCTT	600
Db	3937	ACAGCCACCAAGCAAGCGGAGGCTGCTGCTCCCTGGTAGAATCAAGATGGCGAGCCCTC	3996
OY	601	GAGACCTTCTGGGGGAAACATATGTGAACTTTCAACGGGGAATACAGTACTTATGACAGG	660
Db	3997	GAGACCTTCTGGGGGAAACATATGTGAATTTCAACGGGGAATACATATTTATGACAGG	4056
OY	661	TTGTTCACACTGCTGCGGGAATCCGCGCATTTGCATCATGATGGCGTTCAACAGCTCTGTCT	720
Db	4057	TTGTTCACACTGCTGCGGCAACCCCGCGATAGCATCATGATGGCATTTCAAGCCTCTATCT	4116
OY	721	ACTAGCCGCTACACCCCAATCTACCTCTCTGTTAAATCTTGGGCGGATGGGTAGCC	780
Db	4117	ACGAGCCGCTACACCCCAATACCTCTCTTTTAAATCTTGGGCGGATGGGTAGCC	4176
OY	781	GCCCAACTGCTCCCCCGCATGCTGCTTCAAGCTTTCGATAGGCGCGGATTTGCTGTGGCG	840
Db	4177	GCCCAACTGCTCCCCAGCGCTTCTGCTTTTCGATAGGCGCGGATTTGCTGTGGAGCG	4238

QY	841	GGTGTGGCAGACATAGGCTTGGGAAAGGTGCTTGTCGACATCTTGGCGGGCTAATGAGACA	900
Db	4237	GCTGTGGCAGACATAGGCTTGGGAAAGGTGCTTGTCGATATTTTGGCAGGTTAATGAGACA	4296
QY	901	GGAAGTGGACGCGCGCTGTGGACCTTTAAGGTGCATGACGGCGCAAAATCCCTCCACGAG	960
Db	4297	GGGTGGACGCGCGCTGTGGACCTTTAAGGTGCATGACGGCGCAAAATCCCTCCACGAG	4356
QY	961	GACCTGTAACTTACTCCCTGCCATCTCTCTCTGAGTGCCTGTGTCGGGGTCTGTG	1020
Db	4357	GACCTGTAACTTACTCCCTGCTATCTCTCTCCCTGGCGCCCTAAGTGTGTGGGGTCTGTG	4416
QY	1021	TGCGGAGGAAATCTGCGCTGGCGCATGCGGGTCCAAGGAGGGGGCTGTGCATGATGAAC	1080
Db	4417	TGCGGAGGAAATCTGCGCTGGCGCATGCGGGTCCAAGGAGGGGGCTGTGCATGATGAAC	4476
QY	1081	CGGCTGAATGAGCGTTGCGCTCGGGGGGTAAACATGTTTCCCGCAGCATATATGTGCGAG	1140
Db	4477	CGGCTGAATGAGCGTTGCGCTCGGGGGGTAAACATGTTTCCCGCAGCATATATGTGCTGAG	4536
QY	1141	AGCGACGCGGAGCAGTGTCACTTCAGATCTCTCCGACTTACCTATACCAACATGTG	1200
Db	4537	AGCGACGCTGAGCAGCGTGTCACTTCAGATCTCTCTACCTATACCATCACTCACTGTG	4596
QY	1201	AAGAGGCTTCCACAGTGAATTAAGAGACGTGTCCACGCGCTGTCCGGCTGTGCTA	1260
Db	4597	AAGAGGCTTCCACAGTGAATTAAGAGACGTGTCCACGCGCTGTCCGGCTGTGCTA	4656
QY	1261	AGGATGTTTGGGACGTGATATGACAGTTTTGGTGACTTCAGAACCTGGCTCGAGTC	1320
Db	4657	AGAGATGTTTGGGATTTGATATGACAGGTTTGACTGATTTTCAGAACCTGGCTCGAGTC	4716
QY	1321	AAGCTCTGCGCGCATTAACCGGGAGTCCCTTTTTCATGACCAACGTGGGTACAAGGG	1380
Db	4717	AAGCTCTGCGCGCATTAACCGGGAGTCCCTTTTCATGATCAACGTGGGTACAAGGA	4776
QY	1381	GTTGGCGGGGAGACGGCATATGCAACCACTGTCTATGTGAGACAGATACCGGA	1440
Db	4777	GTTGGCGGGGAGACGGCATATGCAACCACTGTCTATGTGAGACAGATACCGGA	4836
QY	1441	CATGTCAAAAACGTTCCATGATAGATTCGTTGGGCTAAGACTGTAGTAATATGTGGAT	1500
Db	4837	CATGTCAAAAACGTTCCATGATAGATTCGTTGGGCTAAGACTGTAGTAACGTGGAT	4896
QY	1501	GGAACATTTCCCATCAACGATACACACAGGGCCCTCTCAAGCCTTCCCGACGCCAAAC	1560
Db	4897	GGAACATTTCCCATTAACGCGTACACACAGGGCCCTCTCAAGCCTTCCCGACGCCAAAT	4956
QY	1561	TATTCACAGGCGCTGTGGCGGGGTGGCTGTAGAGATACGTGAGATTACGCGGGTGGGG	1620
Db	4957	TATTCACAGGCGCTGTGGCGGGGTGGCTGTAGAGATACGTGAGATTACCGGGTGGGG	5016
QY	1621	GATTTCACATACGTACGACATGACCACTGACCAACGTAAATGCCGTGCCAGATTCCA	1680
Db	5017	GATTTCACATACGTACGACATGACCACTGACCAACGTAAATGCCGTGTCAAGTTCCG	5076
QY	1681	GCCCCCGAATTTCTCACAGAAAGTGAATGGGGTGGCGGCTGCAAGGTAAGCTCCGGCGTGC	1740
Db	5077	GCCCCCGAATTTCTCACAGAAATGGAATGGGGTGGCGGCTGCAAGGTAAGCTTCCAGCTGC	5136
QY	1741	AAACCTCTCTAACGGAGAGAGGTCAATTCACAGGTGGGCTCAACCAATTAACCTGGTTGG	1800
Db	5137	AAACCTCTCTAACGGAGAGAGGTCAATTCCTGAGTGGGCTCAACCAATTAACCTGGTTGG	5196
QY	1801	TGCGAGCTTCCATGCGAGCCGGAACCGGATGTAGAGATGTCTACTTCAATGCTCAACGAC	1860
Db	5197	TGCGAGCTTCCATGTGAGACCCGAACCGGACGTAGAGATGTCTACTTCAATGCTCAACGAC	5256
QY	1861	CCCTCCCATATCAGACAGAGACGGCTTAAGCGCAGAGCTGGCGAGGGGGTCTCCCGCTCC	1920
Db	5257	CCCTCCCATATTCAGCGCGGAGACGGCTTAAGCGTAAAGGTGGCGCGAGGAAATCTCCCGCTCC	5316
QY	1921	TTGGCGACCTCTTACGCTAAGCCAGTGTCTGCGCCCTTCTCGAAGGACATACATTACC	1980

Db	5317	TTGGCCAGCTCATAGCTTAGCCAGCTGTCTGCGCTTCTTTGAAGGCAACTGCACTAC	5376
Qy	1981	CAAAATGACTTCCGAGACGCTGACCTGATCGAGGCCAACTCTGTGGCGGCATGAGATG	2040
Db	5377	CGTCATGACTCCCGGAGCGCTGACCTCATCGAGGCCAACTCTGTGGCGGCAGAGATG	5436
Qy	2041	GGCGGGGACATTACCCCGGTGGAGTCAAGAAACAAGTAGTAGTCTTCTGACTCTTTGAC	2100
Db	5437	GGCGGGAAACAACACCOCGTGGAGTCAAGAAATTAAGTAGTAGTCTTCTGACTCTTTGAC	5496
Qy	2101	CCGCTCCGAGGGGAGAGAGATGACCGGGAAAGTGTCCGTCGCCGGCGGAGATCCTCGGAAA	2160
Db	5497	CCGCTCCAGCGGAGAGAGATGAGAGGAGATATCCGTTCCGCGCGAGATCTGCGGAGG	5556
Qy	2161	TCGAAGAAATTCACACGAGCGATGCCGCATGGGACGCGCGGATTAACAACCTCCGCTG	2220
Db	5557	TCGAAGAAATTCCTCCGAGCGATGCCCATATGGGACGCGCGGATTAACAACCTCCACTG	5616
Qy	2221	CTGAGATCTTGGAAAGCCCGGACTAGTCCCTCCATGTTGATCAATGGTGTCCACTGCA	2280
Db	5617	TTAAGTCTCTGGAAGGACCCCGACTAGCTCCCTCCAGTGTGACACGGGTGTCCATTGGCG	5676
Qy	2281	CCCTCTAAGACCCCTCCATATACACTCCACCGAGGAAGAGACATGTTCTGTACAGAA	2340
Db	5677	CCCTCCAGGCCCCCTCGATACACTCTCCAGGAGGAGAGAGACGTTGTCTCTCAGAA	5736
Qy	2341	TCACACCGTGTCTTCTGCGCTGGCGGAGCTTGCCCAAGAGGCTTTCGGTAGCTCCGAA	2400
Db	5737	TCACACCGTGTCTTCTGCGCTGGCGGAGTGTGCCCAAGAGGCTTTCGGTAGCTCCGAA	5796
Qy	2401	TCGGCCGCTCGACAGCGGACCGGCAACGCGCTCTCTGACCAACCTTCGACGACGGCGGA	2460
Db	5797	TCGGCCGCTCGACAGCGGACCGGCAACGCGCTCTCTGACCAACGCTTCGACGACGGCGAC	5856
Qy	2461	GCAGAGTCTGACGTGAGTGTGATTTCCCTCCAATGCCCGCTTGAAGGGGGAGCGGGGGAC	2520
Db	5857	GCAGAGTCTGACGTGAGTGTGATTTCCCTCCAATGCCCGCTTGAAGGGGGAGCGGGGGAT	5916
Qy	2521	CCCGATCTCAACGACGGGTCTTGGTCTTACCGTGAAGTGAAGAGCGGTGAGGACGTGCTC	2580
Db	5917	CCCGATCTCAACGACGGGTCTTGGTCTTACCGTGAAGTGAAGAGCGGTGAGGACGTGCTC	5976
Qy	2581	TGCTGTCTGATGTCTTACACATGACACAGCGGCTGTGATCACGCCATGCGCTGCGAGGA	2640
Db	5977	TGCTGTCTGATGTCTTACACATGACACAGCGGCTGTGATCACGCCATGCGCTGCGAGGA	6036
Qy	2641	AGCAAGCTGCCCATTCACACGCGTTGAGCACTTTTGTGCTGCTACCAACATAGTCTAC	2700
Db	6037	AGCAAGCTGCCCATTCACACGCGTTGAGCACTTTTGTGCTGCTACCAACATAGTCTAC	6096
Qy	2701	GCTACCAATCCGCGACGCGCAAGCGGAGGAGGAGAAAGTCACTTTTGAACAGCTGCA	2760
Db	6097	GCTACCAATCTCGACGCGCAAGCGCTGGGAGAGAAAGTCACTTTTGAACAGCTGCA	6156
Qy	2761	ATCTGTGACGATCACTACACGAGACGTGTCAAGAGATGAAGGGGAGGCGTCCACATT	2820
Db	6157	GTCTGTGACGACCACTACCGGAGCGTGTCTCAAGAGATGAAGGGGAGGCGTCCACATT	6216
Qy	2821	AAAGCTAAGCTTCTATCACTAGTAGAGAAAGCTGTCAAGCTGACGCCCCCACTTTGGCCAAA	2880
Db	6217	AAAGCTAAGCTTCTATCACTGTGAGAGAAAGCTGTGAAGCTGACGCCCCCACTTTGGCCAGA	6276
Qy	2881	TCTAAATTTGGCTTATGGGGCAAGAGAGCGTCCGAAACCTATCAGACAGGCGCATTTAACAC	2940
Db	6277	TCTAAATTTGGCTTATGGGGCAAGAGAGCGTCCGAAACCTATCAGACAGGCGCATTTAACAC	6336
Qy	2941	ATCCGCTCCGTGTGGAGAGACTTGTGGAGAGACTGAAACACCAATTGACACCAACATC	3000
Db	6337	ATCCGCTCCGTGTGGAGAGACTTGTGGAGAGACTGAGACACCAATTGACACCAACATC	6396
Qy	3001	ATGGCAAAATGAGGTTTTCTGCTGTCAACCAAGAGAGAGGCCCCCAAGCCAGCTGCG	3060

Db 6397 ATGGCAAAAAATGAGTCTTCTGCTCAACAGAGAGGGGGCCGCAAGCCAGCTCGC 6456  
QY 3061 CTTATGCTGTTCCAGACTTGGGGGTCGCTGTCGAGAGAAATGGCCCTCTATGACGTG 3120  
Db 6457 CTTATGCTATTTCCAGATTTGGGGGTTGCTGTGCGAGAAATGGCCCTTTACGATG 6516  
QY 3121 GTCCTCAACCTCCCTCAGGCTGTGATGGGCTCCTCGTACGGAATTCAGATTTCTCTGGA 3180  
Db 6517 GTCTCACACCTCCCTCAGGCGGTGATGGGCTCTTCAATACGGAATTCGAATCTCTCTGGA 6576  
QY 3181 CAGCGGGTCGAGTTCTCTGTGGAACGCTGGAATCAAGAAGACCTCTATGGGCTTTGCA 3240  
Db 6577 CAGCGGGTCGAGTTCTCTGTGGAATGCTGGAAGCGAAGAAATGCCCTATGGGCTTCGCA 6636  
QY 3241 TATGACACCGCGCTGTTTGAAGTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 3300  
Db 6637 TATGACACCGCGCTGTTTGAAGTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 6696  
QY 3301 ATTTATCAATGTTGTGACTTGGCCCCCGAAGCCAGACAGGCAATAGTGTCTCACAGAG 3360  
Db 6697 ATCTACCAATGTTGTGACTTGGCCCCCGAAGCCAGACAGGCAATAGTGTCTCACAGAG 6756  
QY 3361 CGGCTTTATATCGGGGGTCCCTGACTAATTCAAAGGGGAGAACTCGGCTATCCCGG 3420  
Db 6757 CGGCTTTATATCGGGGGTCCCTGACTAATTCAAAGGGGAGAACTCGGCTATCCCGG 6816  
QY 3421 TGCCTGCGGAGCGCGGCTGAGAGTACGCTGCGGCTAATCCCTCAATGTTACTTGAAG 3480  
Db 6817 TGCCTGCGGAGCGCGGCTGAGAGTACGCTGCGGCTAATCCCTCAATGTTACTTGAAG 6876  
QY 3481 GCCTGTGACGCTGTGAGCTGCAAGAGCTCAGAGTCAAGATGCTGTGTGCGAGAG 3540  
Db 6877 GCGGCTGCGGCTGTGAGCTGCAAGAGCTCAGAGTCAAGATGCTGTGTGCGAGAG 6936  
QY 3541 GCGCTGTGCTTATCTGTGAGAGCGCGGGAACCCAGAGAGCGCGGAGACCTTACGAGTC 3600  
Db 6937 GACCTGTGCTTATCTGTGAGAGCGCGGGAACCCAGAGAGCGCGGAGACCTTACGAGTC 6996  
QY 3601 TTACAGGAGCTATGACTAGTACTCTGCCCCCGGGAGCCCGCCCAACAGAAATAC 3660  
Db 6997 TTACAGGAGCTATGACTAGTACTCTGCCCCCGGGAGCCCGCCCAACAGAAATAC 7056  
QY 3661 GACCTGAGTGTGATTAACATCATGCTCTCCATATGTGTGCTGCGGCAAGATGATGCGC 3720  
Db 7057 GACCTGAGTGTGATTAACATCATGCTCTCCATATGTGTGCTGCGGCAAGATGATGCGC 7116  
QY 3721 AAAAGGATATCTAATCTCAACCCGTCGACC 3749  
Db 7117 AAAAGGATATCTAATCTCAACCCGTCGACC 7145

RESULT 5  
AADD5322  
ID AAD25322 standard; cDNA; 7989 BP.  
XX  
AC AAD25322;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Hepatitis C virus (HCV) repliBartMan/Availi cDNA.  
XX  
KM Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
KM gene therapy; vaccine; immunoprotection; hepatotropic; viraemic; liver;  
KM 89.  
XX  
OS Hepatitis C virus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1801..7758  
FT misc\_feature 7766  
FT /\*tag= a  
FT /\*product= "HCVrepliBartMan polypeptide"  
FT /\*tag= b

FT /note= "Nucleotide creating Availi site"  
XX  
PN MO200189364-A2.  
XX  
PD 29-NOV-2001.  
XX  
XX  
PF 23-MAY-2001; 2001WO-US016822.  
XX  
XX 23-MAY-2000; 2000US-00576989.  
XX  
XX (UNITW ) UNIV WASHINGTON.  
XX  
PI Rice CM, Blight KJ;  
XX  
DR MPI; 2002-066755/09.  
DR P-PSDB; ABE15717.  
XX  
PT Hepatitis C virus variants having greater transfection efficiency and  
PT ability to survive subpassage, useful as a vaccine for immunizing primate  
PT to the virus, comprise non-naturally occurring viral sequences.  
XX  
PS Claim 44; Page 69-71; 174pp; English.  
XX  
CC The invention relates to Hepatitis C virus (HCV) variants which include  
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
CC variants that have a transfection efficiency and ability to survive  
CC subpassage greater than HCV that have wild-type polypeptide coding  
CC regions. The polynucleotides of the invention are useful for identifying  
CC a cell line that is permissive for infection with HCV and detecting  
CC replication of HCV in cells of the cell line. They are also useful for  
CC testing a compound for anti-viral properties and for inhibiting HCV  
CC infection. They are also useful for the generation of defined HCV virus  
CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
CC attachment, penetration and entry, structure/function studies on HCV  
CC proteins and RNA elements and identification of new antiviral targets, a  
CC systematic survey of cell culture systems and conditions to identify  
CC those that support wild-type and variant HCV RNA replication and particle  
CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of attenuated or  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilisation of the HCV  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
CC The present sequence is Hepatitis C virus (HCV) repliBartMan/Availi cDNA  
XX  
SQ Sequence 7989 BP; 1647 A; 2369 C; 2242 G; 1731 T; 0 U; 0 Other;  
XX  
Query Match 87.8%; Score 3291.4; DB 6; Length 7989;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;  
QY 1 TGGGAGGGGCTCTTCAAGGCTTCAACGAGTGAATGCCACTTCTGCTCCCAACAAG 60  
Db 3397 TGGGAGGGGCTCTTCAAGGCTTCAACGAGTGAATGCCACTTCTGCTCCCAACAAG 3456  
QY 61 CAGGCAAGAGACAATTTCCCTTACCTGATGCGGTACAGGCTAATGTGTGCGGTAGGGCC 120  
Db 3457 CAGGCAAGAGACAATTTCCCTTACCTGATGCGGTACAGGCTAATGTGTGCGGTAGGGCC 3516  
QY 121 CAGGCCCCACCTTCATCATGAGGATCAAAATGTGGAATGTCTCATAGGCTAAAGCTTACT 180  
Db 3517 CAGGCTCCACCTTCATCATGAGGATCAAAATGTGGAATGTCTCATAGGCTAAAGCTTACT 3576  
QY 181 CAGGCGGGGCAACACCTTGTGTATAGGCTGGGAGCGGTCGCAAGAGGTCACCTTC 240  
Db 3577 CTGCAAGGGCAACGCCCCCTGCTGTATAGGCTGGGAGCGGTCGCAAGAGGTTACTTACC 3636

QY 241 ACAACCCCAATCAAAATTCATATGCGATGTCAGCCGACCTGAGAGTCGTACG 300  
 DB 3637 ACACACCCCATTAACAAATATCATATGCGATGTCAGCTGACCTGGAGGTGTCACG 3696  
 QY 301 AGACCTGGGTGCTGGTGGGCGGGGTCTTGGACGCTCTGGCTGCTATTTGCTTGACACA 360  
 DB 3697 AGACCTGGGTGCTGGTGGGCGGGGTCTTGGACGCTCTGGCTGCTATTTGCTTGACACA 3756  
 QY 361 GGCAGCGTGGTCAATTTGGGTAGATCATCTTGTCCGGGCGGGCGGTATTTGTCGACG 420  
 DB 3757 GGCAGCGTGGTCAATTTGGGTAGATCATCTTGTCCGGGCGGGCGGTATTTGTCGACG 3816  
 QY 421 AGGGAAGTCTCTACCAAGAGATTGATGAGTGAAGAGTGGCGGTGACACTTCCCTTAC 480  
 DB 3817 AGGGAAGTCTCTTACCGGGAGTTGATGAGTGAAGAGTGGCGGTGACACTTCCCTTAC 3876  
 QY 481 ATCGACAGGGAATGCACTGCTGCGGACAGTTCAAGCAAAAAGCGCTCGGTTGTCGAC 540  
 DB 3877 ATCGAAGAGGAATGCACTGCTGCGGACAAATTCAAAACAGAGGCAATCGGTTGTCGAC 3936  
 QY 541 ACAGCCACCAAGAGAGCGGCGGTGCTCCGCTGGTGGAGTCCAGTGGCGAGCCCTT 600  
 DB 3937 ACAGCCACCAAGAGAGCGGCGGTGCTCCGCTGGTGGAGTCCAGTGGCGAGCCCTT 3996  
 QY 601 GAGACCTTCTGGCGGAACAATGTGAACTTTCATGACGCGGATACAGTACTTACGAGG 660  
 DB 3997 GAGACCTTCTGGCGGAACAATGTGAACTTTCATGACGCGGATACAGTACTTACGAGG 4056  
 QY 661 TTGTCACTCTGCTGGGAATCCCGGATGTCATCACTGATGGCGTTCAAGCTCTGTC 720  
 DB 4057 TTGTCACTCTGCTGGGAATCCCGGATGTCATCACTGATGGCGTTCAAGCTCTGTC 4116  
 QY 721 ACTAGCCGCTCAACCAATCTACCTCCGCTTAACTCTGAGGGGAGTGGGTAGG 780  
 DB 4117 ACTAGCCGCTCAACCAATCTACCTCCGCTTAACTCTGAGGGGAGTGGGTAGG 4176  
 QY 781 GCCCACTGCTCCCGCAAGTGTGCTTGAAGTTCGATAGCGCGGCAATGCTGATGCG 840  
 DB 4177 GCCCACTGCTCCCGCAAGTGTGCTTGAAGTTCGATAGCGCGGCAATGCTGATGCG 4236  
 QY 841 GCTGTTGCGACATAGGCTTTGGGAAGTGTGGAACATCTTGGCGGGCTATGAGCA 900  
 DB 4237 GCTGTTGCGACATAGGCTTTGGGAAGTGTGGAACATCTTGGCGGGCTATGAGCA 4296  
 QY 901 GGAAGTGGCAGGCGCGCTGCTGGCTTTAAGTTCATAGAGGCGGAATGCCCCCTCACCGAG 960  
 DB 4297 GGAAGTGGCAGGCGCGCTGCTGGCTTTAAGTTCATAGAGGCGGAATGCCCCCTCACCGAG 4356  
 QY 961 GACTGTTAACTTACCTGCTGATCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
 DB 4357 GACTGTTAACTTACCTGCTGATCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 4416  
 QY 1021 TSCGACGAGTACTGCTGCGGACGCTGAGTTCAGAGGAGGAGGCTTGGCATGATGAAC 1080  
 DB 4417 TSCGACGAGTACTGCTGCGGACGCTGAGTTCAGAGGAGGAGGCTTGGCATGATGAAC 4476  
 QY 1081 CGGCTGATAGCTTGGCTGCGGGGTAAACATGTTTCCCGCAAGCATATGTCGACAG 1140  
 DB 4477 CGGCTGATAGCTTGGCTGCGGGGTAAACATGTTTCCCGCAAGCATATGTCGACAG 4536  
 QY 1141 AGCGAGCGCGAGCATGCTGATCTCTGACCTTACCTTACCTTACCTTACCTTACCTT 1200  
 DB 4537 AGCGAGCGCGAGCATGCTGATCTCTGACCTTACCTTACCTTACCTTACCTTACCTT 4596  
 QY 1201 AAGAGCTTCACAGTGAATTAACAGAGCTGCTCAGCGCCCTGCTCGGCTGTCGCTA 1260  
 DB 4597 AAGAGCTTCACAGTGAATTAACAGAGCTGCTCAGCGCCCTGCTCGGCTGTCGCTA 4656  
 QY 1261 AAGAGTCTTGGGACGATGATGACAGTGGTGGCTGACTTAAAGACTGGCTCCAGTCC 1320  
 DB 4657 AAGAGTCTTGGGACGATGATGACAGTGGTGGCTGACTTAAAGACTGGCTCCAGTCC 4716  
 QY 1321 AAGCTCTGCGGATTAACGGGAGTCCCTTTTCTCATGCAAGCTGAGGATCAAGGGG 1380

DB 4717 AAGCTCTGCGGATTTGCGGGAGTCCCTTCTTCTCATGTCACAGTGGATCAAGGGG 4776  
 QY 1381 GTCTGGCGGGGAGACGGCATCATGACAGACCACTGCTCATGTGAGACACAGATCACCGA 1440  
 DB 4777 GTCTGGCGGGGAGACGGCATCATGACAGACCACTGCTCATGTGAGACACAGATCACCGA 4836  
 QY 1441 CATGTCAAAAACGGTTCCATGAGGATCGTTGGGCTTAAACCTGTAGTAACTGTGGCAT 1500  
 DB 4837 CATGTCAAAAACGGTTCCATGAGGATCGTTGGGCTTAAACCTGTAGTAACTGTGGCAT 4896  
 QY 1501 GGAACATTCGCCATCAACGCATACACAGGAGCCCTGACGCGCTCCCAAGGCCAAAC 1560  
 DB 4897 GGAACATTCGCCATTAACGCATACACAGGAGCCCTGACGCGCTCCCAAGGCCAAAT 4956  
 QY 1561 TATTCCAGGCGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGATTACCGGGGTGGG 1620  
 DB 4957 TATTCCAGGCGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGATTACCGGGGTGGG 5016  
 QY 1621 GATTTCACATGATGACGAGCATGACCACTGACCAAGTAAATGCCCCGAGGTTCCA 1680  
 DB 5017 GATTTCACATGATGACGAGCATGACCACTGACCAAGTAAATGCCCCGAGGTTCCA 5076  
 QY 1681 GCCCCGGAATCTTCAACAGAGTGAAGTGGGTGCGGCTGACAGGATACGCTCCGCGTGC 1740  
 DB 5077 GCCCCGGAATCTTCAACAGAGTGAAGTGGGTGCGGCTGACAGGATACGCTCCGCGTGC 5136  
 QY 1741 AAACCTCTCTACGGAAGAGGTCACTTTCAGGTGCGGCTCAACCAATACCTGTTGGG 1800  
 DB 5137 AAACCTCTCTACGGAAGAGGTCACTTTCAGGTGCGGCTCAACCAATACCTGTTGGG 5196  
 QY 1801 TGGCAGCTCCATGGAAGCGGACCGGATGAGAGTGTCACTTCCATGCTCACCGAC 1860  
 DB 5197 TGGCAGCTCCATGGAAGCGGACCGGATGAGAGTGTCACTTCCATGCTCACCGAC 5256  
 QY 1861 CCTCCCATCATCAACAGAGAGCGGCTAGCGCAGGCTGGCCAGGAGGCTCTCCCTCC 1920  
 DB 5257 CCTCCCATCATCAACAGAGAGCGGCTAGCGCAGGCTGGCCAGGAGGCTCTCCCTCC 5316  
 QY 1921 TTGGCAGCTCTTACAGTACGCCAGTTGTGTGCGCTTCTCGAAGGCGCATATATAC 1980  
 DB 5317 TTGGCAGCTCTTACAGTACGCCAGTTGTGTGCGCTTCTCGAAGGCGCATATATAC 5376  
 QY 1981 CAAATGACTTCCGAGCGCTGACCTCATTCGAGGCGCAACCTCTGCGGGCATAGAG 2040  
 DB 5377 CAAATGACTTCCGAGCGCTGACCTCATTCGAGGCGCAACCTCTGCGGGCATAGAG 5436  
 QY 2041 GGCAGGGAATTAACCGCGTGAAGTCAAGAAACAAGTATGATCTGGAATCTTTGAC 2100  
 DB 5437 GGCAGGGAATTAACCGCGTGAAGTCAAGAAACAAGTATGATCTTTGGAATCTTTGAC 5496  
 QY 2101 CCGCTCCGAGCGAGAGATGAGCGGAAAGTGTCCGTCGCGCGAGATCTTCGCGAAA 2160  
 DB 5497 CCGCTCCGAGCGAGAGATGAGCGGAAAGTGTCCGTCGCGCGAGATCTTCGCGAAA 5556  
 QY 2161 TCCAGAAATTTCCACAGCGATGCGCGCATGCGGATGCGGAGTTTAAACCTCCGCTG 2220  
 DB 5557 TCCAGAAATTTCCACAGCGATGCGCGCATGCGGATGCGGAGTTTAAACCTCCGCTG 5616  
 QY 2221 CTGGAAGTCTGGAAGGCGCGGACTACGCTCTCAGTGGTACATGAGGCTCCACATGCA 2280  
 DB 5617 CTGGAAGTCTGGAAGGCGCGGACTACGCTCTCAGTGGTACATGAGGCTCCACATGCA 5676  
 QY 2281 CCTACTAAGACCTCTCTATATACCTTCAACGAGAGAAAGAGACAGTTGTTGACAGAA 2340  
 DB 5677 CCTACTAAGACCTCTCTATATACCTTCAACGAGAGAAAGAGACAGTTGTTGACAGAA 5736  
 QY 2341 TCCACCGGTCTTTCGCGCGGAGGCTGCGCAAAAGCTTTCGAGTACGCTCCGAAACG 2400  
 DB 5737 TCCACCGGTCTTTCGCGCGGAGGCTGCGCAAAAGCTTTCGAGTACGCTCCGAAACG 5796  
 QY 2401 TCGGCGTTCAGAGGAGACGGCAACCGCCCTCTGACCAACCTCCGACGACGCGGAG 2460

Db 5797 TCAGCGCTGACAGCGGCAACGCAACGCGCTCTCTGACCAAGCCCTCCGACGACGCGAC 5856  
 Qy 2461 GAGAGATCGACGCTTGAAGTGTATTCCTCCATGCCCCCTTGAAGGGGGAGCGCGGGGAC 2520  
 Db 5857 GCGGGATCCGACGCTGATGTACTCTCCATGCCCCCTTGAAGGGGGAGCGCGGGGAT 5916  
 Qy 2521 CCCGATCTCAGCAGCGGGTCTTGGTCTACCGTGAAGTGAAGGCGGGTGAAGAGAGTCTG 2580  
 Db 5917 CCGGATCTCAGCAGCGGGTCTTGGTCTACCGTGAAGGAGGAGGCTAGTGAAGAGTCTG 5976  
 Qy 2581 TGGTGTCTGATGTCTTACATGATGAGAGCGGCTCTGATACCGCATGCGCTCGAGAGAA 2640  
 Db 5977 TGGTGTCTGATGTCTTACATGATGAGAGCGGCTCTGATACCGCATGCGCTCGAGAGAA 6036  
 Qy 2641 AGCAAGCTGCGCATCAACGCGTGAAGCACTCTTGTGTGATCAACAAATGATGTCTAC 2700  
 Db 6037 ACCAAGCTGCGCATCAATGCACTGAGCACTTGTGTGATCAACAAATGATGTCTAC 6096  
 Qy 2701 GCTACCAATCCGCGACGCGCAACGCGGACGAGAGAGTCACTTTTGAACAGACTGCA 2760  
 Db 6097 GCTACCAATCTGCGACGCGCAACGCTGCGGAGAGAGTCACTTTTGAACAGACTGCA 6156  
 Qy 2761 ATCTTGAGAGATCACTACAGAGAGTGTCTCAAGAGATGAAGGCGGAGCGCTCAAGATT 2820  
 Db 6157 GTCTGAGAGAGATCACTACAGAGAGTGTCTCAAGAGATGAAGGCGGAGCGCTCAAGATT 6216  
 Qy 2821 AAGGCTTAAGCTTATCTACAGAGAGAGAGCTGCAAGCTGAGCGCCCACTTGGGCGCAA 2880  
 Db 6217 AAGGCTTAAGCTTATCTACAGAGAGAGAGCTGTAAGTGAAGCGCCCACTTGGGCGCAA 6276  
 Qy 2881 TCTAAATTTGGCTATGAGGGGCAAGAGAGTCCGGAACCTATCAGAGAGCGCATTTAACAC 2940  
 Db 6277 TCTAAATTTGGCTATGAGGGGCAAGAGAGTCCGGAACCTATCAGAGAGCGCATTTAACAC 6336  
 Qy 2941 ATCCGCTCCGTGTGAGAGAGCTTGTGGAAGACCTGAACACCAATTTGACACCAACATC 3000  
 Db 6337 ATCCGCTCCGTGTGAGAGAGCTTGTGGAAGAGACCTGAACACCAATTTGACACCAACATC 6396  
 Qy 3001 ATGGCAAAAATAGAGGTTTCTGCTCAACCAAGAGAGAGGCGGCAAGCGACGCTGCG 3060  
 Db 6397 ATGGCAAAAATAGAGGTTTCTGCTCAACCAAGAGAGAGGCGGCGCAAGCGACGCTGCG 6456  
 Qy 3061 CTATCGTGTCCAGACTTGGGGGTCCTGTGTGAGAGAGAGAGGCGCTCTATGAGAGT 3120  
 Db 6457 CTATCGTGTCCAGACTTGGGGGTCCTGTGTGAGAGAGAGAGGCGCTCTATGAGAGT 6516  
 Qy 3121 GTCTCAACCTCCCTCAGGCTGTGATGGGCTCTCTGTAAGGATTTCAATTTCTCTGGA 3180  
 Db 6517 GTCTCAACCTCCCTCAGGCTGTGATGGGCTCTCTGTAAGGATTTCAATTTCTCTGGA 6576  
 Qy 3181 CAGCGGTTGAGTTCCTGTGGAAGCGCTGGAATCAAGAGAGAGCGCTATGGGCTTTGCA 3240  
 Db 6577 CAGCGGTTGAGTTCCTGTGGAAGCGCTGGAAGAGAGAGAGAGCGCTATGGGCTTTGCA 6636  
 Qy 3241 TATGACACCGGCTTTTGAATCAACAGTCACTGAGATGACATCCGTGTGAGAGAGTCA 3300  
 Db 6637 TATGACACCGGCTTTTGAATCAACAGTCACTGAGATGACATCCGTGTGAGAGAGTCA 6696  
 Qy 3301 ATTATCAATGTGTGACTTGGCCCCCGAAGCAGACAGCGCATAGGCTGCTCACAGAG 3360  
 Db 6697 ATTATCAATGTGTGACTTGGCCCCCGAAGCAGACAGCGCATAGGCTGCTCACAGAG 6756  
 Qy 3361 CCGGTTTATATTCGGGGGTTCCCTGACATTAATTTCAAAAGGAGAGAGTGGGCTATCCGCGG 3420  
 Db 6757 CCGGTTTATATTCGGGGGTTCCCTGACATTAATTTCAAAAGGAGAGAGTGGGCTATCCGCGG 6816  
 Qy 3421 TGGCGCGGAGCGGCTGTGACGACTAGCTGCGGTATATCCCTCAAGATTTACTTTGAG 3480  
 Db 6817 TGGCGCGGAGCGGCTGTGACGACTAGCTGCGGTATATCCCTCAAGATTTACTTTGAG 6876  
 Qy 3481 GCCCTTGACGCTGTGAGCTGCAAGCTTCAGAGACTGACGATGCTGTGTGCGAGAC 3540  
 Db 6877 GCCCTTGACGCTGTGAGCTGCAAGCTTCAGAGACTGACGATGCTGTGTGCGAGAC 6936

Qy 3541 GGCCTTGCTATCTGTGAGAGCGCGGAGACCCAGAGAGCGCGGAGCTTACGAGTTC 3600  
 Db 6937 GACCTTGCTATCTGTGAAAGCGCGGAGACCCAGAGAGAGAGCGGAGCTTACGAGTTC 6996  
 Qy 3601 TTACGAGAGGCTATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3660  
 Db 6997 TTACGAGAGGCTATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7056  
 Qy 3661 GACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720  
 Db 7057 GACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7116  
 Qy 3721 AAAAGGTTACTACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3749  
 Db 7117 AAAAGGTTACTACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 7145

RESULT 6  
 ADJ57845  
 ID ADJ57845 standard; DNA; 7989 BP.  
 XX  
 AC ADJ57845;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DB HCV replicon encoding sequence.  
 XX  
 KW hepatitis C virus; HCV; Antinflammatory; Hepatotropic; Virucide; de;  
 KM HCV replicon.  
 XX  
 OS Unidentified.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 1801..7759  
 FT FT /tag= a  
 FT FT /product= "HCV replicon"  
 XX  
 PN MO2004015131-AZ.  
 XX  
 PD 19-FEB-2004.  
 XX  
 PF 12-AUG-2003; 2003WO-US025260.  
 XX  
 PR 12-AUG-2002; 2002US-040261P.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Gao M, Lemm JA, O'Boyle DR, Nower P, Rigat K, Sun J;  
 XX  
 DR WPI; 2004-180685/17.  
 DR P-PSDB; ADJ57846.  
 XX  
 PT Use of hepatitis C virus assays or reporter assays, e.g. identifying a  
 PT compound that inhibits hepatitis C virus RNA replication or identifying a  
 XX compound that modulates the activity of a gene of interest.  
 PS Claim 3; SEQ ID NO 1; 45pp; English.  
 XX  
 CC The present invention relates to the use of hepatitis C virus (HCV)  
 CC assays for identifying a compound that inhibits HCV RNA replication and  
 CC reporter assays for identifying a compound that modulates the activity of  
 CC a gene of interest. The assays are useful for identifying a compound that  
 CC inhibits HCV RNA replication or for identifying a compound that modulates  
 CC the activity of a gene of interest. The HCV assay is useful for high  
 CC throughput screening that quantifies both the amount of HCV RNA  
 CC replication inhibitory activity associated with a test compound and the  
 CC amount of cytotoxicity associated with the test compound. The compound is  
 CC useful for treating hepatitis C infection. Assays of the invention have  
 CC distinct advantages when compared to qRT-PCR or other methods in that  
 CC assays of the invention may take place in situ in a detergent based crude  
 CC cell lysate, which requires no further preparation prior to performing  
 CC the assays. The assays do not also involve numerous manipulations to add



CC or subtract reagents after addition of test compounds and are desirably based on a viral protein which is required by the HCV replicon for CC replication. The present sequence represents a HCV replicon encoding CC sequence used in the assay of the invention.

XX Sequence 7989 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other;

Query Match 87.8%; Score 3291.4; DB 12; Length 7989;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

```

QY 1 TGGGAGGCGCTCTTCAAGGCTCAGCCAGGTGAGTCCCACTTCTGTCCTCAAG 60
DB 3397 TGGGAGGCGCTCTTCAAGGCTCAGCCAGGTGAGTCCCACTTCTGTCCTCAAG 3456
QY 61 CAGGAGAGAGACAATTTCCCTTACCTGAGGCTTACAGGCTACTGTGTGCGTACGG 120
DB 3457 CAGGAGAGAGACAATTTCCCTTACCTGAGGCTTACAGGCTACTGTGTGCGTACGG 3516
QY 121 CAGGAGGCGCTCTTCAAGGCTCAGCCAGGTGAGTCCCACTTCTGTCCTCAAG 180
DB 3517 CAGGAGGCGCTCTTCAAGGCTCAGCCAGGTGAGTCCCACTTCTGTCCTCAAG 3576
QY 181 CTGCGGCGGCGCAACCCCTTGTATAGGCTGAGGCGGCTCCGAAACGAGTCAAC 240
DB 3577 CTGCGGCGGCGCAACCCCTTGTATAGGCTGAGGCGGCTCCGAAACGAGTCAAC 3636
QY 241 ACACACCCCATTAACCAATTCATGATGAGTGCATGTCAGCCGACTGAGGCTG 300
DB 3637 ACACACCCCATTAACCAATTCATGATGAGTGCATGTCAGCCGACTGAGGCTG 3696
QY 301 AGACCTGGGTGTGAGGCGGCGGCTTGCAGCTTGTGCTGCTGCTGCTGCTG 360
DB 3697 AGACCTGGGTGTGAGGCGGCGGCTTGCAGCTTGTGCTGCTGCTGCTGCTG 3756
QY 361 GGCAGCGTGTCTATGTTGAGTATGATCTTGTCCGCGCGCGGCTATTTGTTCC 420
DB 3757 GGCAGCGTGTCTATGTTGAGTATGATCTTGTCCGCGCGCGGCTATTTGTTCC 3816
QY 421 AGGGAAGTCTCTTACCAAGATTTGATGATGAGAGAGTGCCTGCTGCTGCTTAC 480
DB 3817 AGGGAAGTCTCTTACCAAGATTTGATGATGAGAGAGTGCCTGCTGCTGCTTAC 3876
QY 481 ATGACAGAGGAGATGAGCTGCGGAGAGTTCAGGAAAGAGCGCTCGGCTGCTG 540
DB 3877 ATGACAGAGGAGATGAGCTGCGGAGAGTTCAGGAAAGAGCGCTCGGCTGCTG 3936
QY 541 ACAGCACAAGAGAGAGGCGGCTGCTCCGTTGAGTCCAAAGTGGCGAGCCCTT 600
DB 3937 ACAGCACAAGAGAGAGGCGGCTGCTCCGTTGAGTCCAAAGTGGCGAGCCCTT 3996
QY 601 GAGACCTTCTGGGCGAAACATGATGAGACTTCATCAGCGGAGTACAGTACG 660
DB 3997 GAGACCTTCTGGGCGAAACATGATGAGACTTCATCAGCGGAGTACAGTACG 4056
QY 661 TTGTCACTCTGCTGGGAAATCCGCGATTTGATGATGAGTGCCTGCTGCTGCT 720
DB 4057 TTGTCACTCTGCTGGGAAATCCGCGATTTGATGATGAGTGCCTGCTGCTGCT 4116
QY 721 ACTAGCCGCTCAACCAATCTACCTCTGCTTAACTCTGCGGAGGATGGGTAG 780
DB 4117 ACTAGCCGCTCAACCAATCTACCTCTGCTTAACTCTGCGGAGGATGGGTAG 4176
QY 781 GCCCAACTGCTCCCCAGAGTGTCTTCACTTGTGAGGCGCGGATGTGCTGCTG 840
DB 4177 GCCCAACTGCTCCCCAGAGTGTCTTCACTTGTGAGGCGCGGATGTGCTGCTG 4236
QY 841 GCTGTTGAGAGATAGGCTTTGGGAAGTCTTGTGATCTTTGGGCGGCTATGAGCA 900
DB 4237 GCTGTTGAGAGATAGGCTTTGGGAAGTCTTGTGATCTTTGGGCGGCTATGAGCA 4296
QY 901 GGAGTGGAGCGCGCTGCTGAGCTTTAAGTATGAGCGGCGAAATGCTCCACGAG 960

```

```

DB 4297 GGGGTGAGGCGCGCTGCTGCTTAAAGTATGAGCGGCGAGATGCCCTCCACGAG 4356
QY 961 GACCTGGTTAACTTACCTCCGTCATCTCTCTGTCGTCGTCGTCGTCGTCGTCG 1020
DB 4357 GACCTGGTTAACTTACCTCCGTCATCTCTCTGTCGTCGTCGTCGTCGTCGTCG 4416
QY 1021 TGGCAGCGATCTGCTGCGGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1080
DB 4417 TGGCAGCGATCTGCTGCGGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4476
QY 1081 CGGCTAATAGCTTCCCTCTGCGGCGTAAACATGTTTCCCGACGATATGTCAGAG 1140
DB 4477 CGGCTAATAGCTTCCCTCTGCGGCGTAAACATGTTTCCCGACGATATGTCAGAG 4536
QY 1141 AGGAGCGCGACAGTGTCACTAGATCTCTCCGACCTTACTATACCCGAACTGTTG 1200
DB 4537 AGGAGCGCGACAGTGTCACTAGATCTCTCCGACCTTACTATACCCGAACTGTTG 4596
QY 1201 AAGAGCTCCACAGTGAATTAACGAGACTGCTCCAGCGCTGCTCCGCTGTCGCTA 1260
DB 4597 AAGAGCTCCACAGTGAATTAACGAGACTGCTCCAGCGCTGCTCCGCTGTCGCTA 4656
QY 1261 AAGGATGTTTGGACTGATATGACAGATTTTGGCTGACTTCAAGACTGCTCCAGTCC 1320
DB 4657 AAGGATGTTTGGACTGATATGACAGATTTTGGCTGACTTCAAGACTGCTCCAGTCC 4716
QY 1321 AAGCTCTGCGGAGTATCCGGAATCCCTTTTCTCATGCTCCAAAGTGGTCAAGGGG 1380
DB 4717 AAGCTCTGCGGAGTATCCGGAATCCCTTTTCTCATGCTCCAAAGTGGTCAAGGGG 4776
QY 1381 GTCTGCGGAGAGACGAGCATCATGACAGCAACCTGCTCATGTCAGTGAAGACATCACGGA 1440
DB 4777 GTCTGCGGAGAGACGAGCATCATGACAGCAACCTGCTCATGTCAGTGAAGACATCACGGA 4836
QY 1441 CATGTCAAAAACGTTTCATAGAGATGTTTGGCTTAAGACTGTATGAATCATGTGCGAT 1500
DB 4837 CATGTCAAAAACGTTTCATAGAGATGTTTGGCTTAAGACTGTATGAATCATGTGCGAT 4896
QY 1501 GGAACATTTCCCATTAACGATTAACAGAGGCGGCTTGAACGCGCTCCGAGGCGCAAC 1560
DB 4897 GGAACATTTCCCATTAACGATTAACAGAGGCGGCTTGAACGCGCTCCGAGGCGCAAC 4956
QY 1561 TATTCAAGGCGCTGTCGCGGCTGCTGCTGAGAGTACGTCGAGGTTACGCGGCTGAGG 1620
DB 4957 TATTCAAGGCGCTGTCGCGGCTGCTGCTGAGAGTACGTCGAGGTTACGCGGCTGAGG 5016
QY 1621 GATTTCACATGCTGAAGAGATGACACTGACCAAGTAAATGCGCTGCGAGTTCCA 1680
DB 5017 GATTTCACATGCTGAAGAGATGACACTGACCAAGTAAATGCGCTGCGAGTTCCG 5076
QY 1681 GCGCCCGAATTTCTTCAAGAGTGGATGGGCTGCTGCAAGTACCTCCGCGGTGC 1740
DB 5077 GCGCCCGAATTTCTTCAAGAGTGGATGGGCTGCTGCAAGTACCTCCGCGGTGC 5136
QY 1741 AAACCTCTCTTACGAGAGAGTCACTTCCAGGCTGAGGCTCAACCAATGCTGTTGG 1800
DB 5137 AAACCTCTCTTACGAGAGAGTCACTTCCAGGCTGAGGCTCAACCAATGCTGTTGG 5196
QY 1801 TCGAGCTCCATGCGAGCGGCGAACCGAGTGTAGAGTGTCTCAATGCTCACGAG 1860
DB 5197 TCGAGCTCCATGCGAGCGGCGAACCGAGTGTAGAGTGTCTCAATGCTCACGAG 5256
QY 1861 CCTTCCACATCAGAGAGAGAGCTTAAAGCGAGCTGCGCAGGCGGCTTCCCTCC 1920
DB 5257 CCTTCCACATCAGAGAGAGAGCTTAAAGCGAGCTGCGCAGGCGGCTTCCCTCC 5316
QY 1921 TTGGCAGGCTTCACTAGCGAGTGTGTCGCGCTTCTCCGAAAGCGCATCATTTAC 1980
DB 5317 TTGGCAGGCTTCACTAGCGAGTGTGTCGCGCTTCTCCGAAAGCGCATCATTTAC 5376
QY 1981 CAAATGACTTCCACAGCTGACCTCATGAGGCCAATCTCTGTGTGGCGCATGAGATG 2040
DB 5377 CGTATGACTTCCCGAGCGGTGACCTCATGAGGCCAATCTCTGTGTGGCGCATGAGATG 5436

```

QY 2041 GGGGGGAGATTACCCCGGTGAGTCAGAGAACAGTAGTATCCCTGACCTTTTCAGC 2100  
 DB 5437 GGGGGGAACATCCCGCGGTGAGTCAGAGAAATAGGTATATTTTGGACTTTTCGAG 5496  
 QY 2101 CCGCTCCGAGCGGAGAGATGAGCGGAAATGTCCGTCCGGCGGAGATCTCGCGAAA 2160  
 DB 5497 CCGCTCCAGCGGAGAGATGAGAGGAATATCCGTTCCGGCGGAGATCCCGCGAGG 5556  
 QY 2161 TCCAAAGAAATTCACAGAGATGCGCGCATGAGGAGCGCCGGATTACAACTCCGCTG 2220  
 DB 5557 TCCAGAAATTCCTGAGGAGATGCCATATGAGGACAGCCGGATTAACAACCTTCACTG 5616  
 QY 2221 CTGAGATCTGAGAGGCGCCGAGCTACGTCCCTCAAGTGTAATGATGTCCTCCCACTGCA 2280  
 DB 5617 TTAGAGTCTGAGAGGACCCGAGCTAGTCCCTCCAGTGTATACGGGTGTCCATTGGCGG 5676  
 QY 2281 CCTACTAAGACCCCTCTTATATCACTCCAGAGAGAGAGACAAGTTGTTGTGACAGAA 2340  
 DB 5677 CTGCAAGGCGCCCTCCGATACCACTCCAGAGAGAGAGACGGTTGTCTGTGAGAA 5736  
 QY 2341 TCCACCGTGTCTCTGCGGTGAGAGCTTCCCAAGGCTTTCGGTGTCCGAAACCG 2400  
 DB 5737 TCTACCGTGTCTTGTGCTTGGCGAGCTGCGCAAAAGACCTTCGAGCTCCGAAATCG 5796  
 QY 2401 TCGGCGGTGACAGCGGACAGGCAACGCGCTCTGACCAACCTTCGAGAGCGGCGGA 2460  
 DB 5797 TCGGCGGTGACAGCGGACAGGCAACGCGCTCTGACCAACGCGCTTCGAGAGCGGCGAC 5856  
 QY 2461 GCAGATCTGACGTTGATGTCATATTCCTCATATGCCCCCTTTAGAGGGAGCGGAGGAC 2520  
 DB 5857 GCGGAGTCCGACGTTGATGTCATCTCTCATATGCCCCCTTTAGAGGGAGCGGAGGAT 5916  
 QY 2521 CCGGATCTGACGAGCGGTCTTGTCTCATCGTGAAGAGAGCGCGTGTGAGAGCTGTC 2580  
 DB 5917 CCGGATCTGACGAGCGGTCTTGTCTCATCGTGAAGAGAGCGCTGATGAGAGCTGTC 5976  
 QY 2581 TGTGTCGATGTCCTACATGAGAGAGCGCTCTGATCAAGCCATCGCTGCGAGGAA 2640  
 DB 5977 TGTGTCGATGTCCTACATGAGAGAGCGCGCTGATCAAGCCATCGCTGCGAGGAA 6036  
 QY 2641 AGCAAGCTGCCATCAACGCGTTGAGCACTCTTGTGCTGCAACAATGCTGTAC 2700  
 DB 6037 ACCAAGCTGCCATCAATGACTGAGCACTCTTGTGCTGCAACAATGCTGTAT 6096  
 QY 2701 GCTACCAATCCCGAGCGGAGAGCGGACAGAGAGGTCCTTTGACAGATGCA 2760  
 DB 6097 GCTACCAATCCCGAGCGGAGAGCGGACAGAGAGGTCACCTTTGACAGATGCA 6156  
 QY 2761 ATCTGAGCATCACTACAGAGAGTGTCTCAAGAGATGAGAGGAGGTCACAGTT 2820  
 DB 6157 GTCTGAGCATCACTACCGGAGAGTGTCTCAAGAGATGAGAGGAGGTCACAGTT 6216  
 QY 2821 AAGGCTAAGCTTCTATCATGTAAGAGAGCTGCAAGTGAACGCCCCCACTTCGCGCAA 2880  
 DB 6217 AAGGCTAAGCTTCTATCCGTGAGAGAGCTGTAGCTGACGCCCCCACTTCGCGCAGA 6276  
 QY 2881 TCTAAATTTGGCTATGAGGAGAGAGAGAGTCCGAACTTATCCAGAGAGGCTTAAACAC 2940  
 DB 6277 TCTAAATTTGGCTATGAGGAGAGAGAGAGTCCGAACTTATCCAGAGAGGCTTAAACAC 6336  
 QY 2941 ATCCGCTCCGTGAGAGAGCTTGTGTAAGAGACTGAAACACCAATTGACACCAATC 3000  
 DB 6337 ATCCGCTCCGTGAGAGAGCTTGTGTAAGAGACTGAAACACCAATTGACACCAATC 6396  
 QY 3001 ATGCAAAAATAGAGGTTTTTCTGCTCAACAGAGAGAGAGGCGCAAGCCAGCTGCG 3060  
 DB 6397 ATGCAAAAATAGAGGTTTTTCTGCTCAACAGAGAGAGAGGCGCAAGCCAGCTGCG 6456  
 QY 3061 CTTATCGTGTCCAGACTTGGGGGTCCGTGTGTCAGAGAAATGAGCCCTTATGAGTG 3120  
 DB 6457 CTTATCGTGTCCAGACTTGGGGGTCCGTGTGTCAGAGAAATGAGCCCTTATGAGTG 6516

QY 3121 GTCTCACCTCTCCCTCAGGCTGTGATGAGGCTCTCTGTACGATTCAGATTTCTCTGGA 3180  
 DB 6517 GTCTCACCTCTCCCTCAGGCTGTGATGAGGCTCTCTGTATACGGAATTCATTAATCTCTGGA 6576  
 QY 3181 CAGCGGTGAGATTCCTGTGTAAGCGCTGGAATCAAGAGAAACCCCTATGAGGCTTTGCA 3240  
 DB 6577 CAGCGGTGAGATTCCTGTGTAAGCGCTGGAATCAAGAGAAATGCCCTATGAGGCTTTGCA 6636  
 QY 3241 TATGACACCGCTGTTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 3300  
 DB 6637 TATGACACCGCTGTTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 6696  
 QY 3301 ATTATCAATGTTGTGACTTGGCCCCCGAAGCCAGACAGGCAATAGTGTCTCACAGG 3360  
 DB 6697 ATCTACAAATGTTGTGACTTGGCCCCCGAAGCCAGACAGGCAATAGTGTCTCACAGG 6756  
 QY 3361 CCGCTTATATGAGGAGTCCCTGACTAATTCAAAGGCGAGAACTGGGGCTATGCGCGG 3420  
 DB 6757 CCGCTTATATGAGGAGTCCCTGACTAATTCAAAGGCGAGAACTGGGGCTATGCGCGG 6816  
 QY 3421 TGCAGGCGAGGCGGTGCTGACGACTAGTGCAGTATATACCTTCATATGTTAAG 3480  
 DB 6817 TGCAGGCGAGGCGGTGCTGACGACTAGTGCAGTATATACCTTCATATGTTAAG 6876  
 QY 3481 GCTCTGACGCTGTGACCTGCAAACTGTCAGAGACTGACGATGCTGTGCGGAGAC 3540  
 DB 6877 GCGGTGCGGCTGTGACCTGCAAACTGTCAGAGACTGACGATGCTGTGCGGAGAC 6936  
 QY 3541 GCGCTGTGATATCTGTGAGAGCGCGGAGAACCCAGAGAGAGCGGCGGAGCTTACGAGAC 3600  
 DB 6937 GACCTGTGATATCTGTGAGAGCGCGGAGAACCCAGAGAGAGCGGCGGAGCTTACGAGAC 6996  
 QY 3601 TTCAGGAGGCTATGACTAGTACTGCCCCCGGAGGAGCCCGCCCAACAGAAATAC 3660  
 DB 6997 TTCAGGAGGCTATGACTAGTACTGCCCCCGGAGGAGCCCGCCCAACAGAAATAC 7056  
 QY 3661 GACCTGAGTATTAATCATATGCTCTTCAATGTTGTGAGTGTGCGGACGATGATGTCG 3720  
 DB 7057 GACCTGAGTATTAATCATATGCTCTTCAATGTTGTGAGTGTGCGGACGATGATGTCG 7116  
 QY 3721 AAAAGGTATTAATCACTCAACCGTGACCC 3749  
 DB 7117 AAAAGGTATTAATCACTCAACCGTGACCC 7145

RESULT 7  
 AAL47276  
 ID AAL47276 standard; DNA; 7992 BP.  
 XX  
 AC AAL47276;  
 AC  
 DT 30-AUG-2002 (first entry)  
 XX  
 DE Hepatitis C virus sub-genomic replicon clone I377-NS3-3'UTR.  
 XX  
 KM Hepatitis; HCV; core-neo; NS3 proteinase/helicase; vaccine; diagnosis;  
 KM virulence; hepatotropic; gene therapy; anti-viral; gene; de.  
 OS Hepatitis C virus.  
 XX  
 XX  
 FH Key  
 FT CDS 342..1181 Location/Qualifiers  
 FT /\*tag= a  
 FT /product= "core-neo fusion protein"  
 FT 1801..7758  
 FT /\*tag= b  
 FT /product= "NS3 proteinase/helicase"  
 PN MO200238793-A2.  
 XX  
 PD 16-MAY-2002.  
 XX  
 PF 02-NOV-2001; 2001MO-US046350.



XX 07-NOV-2000; 2000US-0245866P.  
 PR (ANAD-) ANADYS PHARM INC.  
 PA Blichko V;  
 XX WPI; 2002-490082/52.  
 DR P-PSDB; AAO18000, AAO18001.  
 XX  
 PT Novel nucleic acid encoding replication competent recombinant hepatitis C  
 PT virus genome useful for screening anti-hepatitis C virus therapeutics and  
 PT for vaccine development.  
 PS  
 XX Claim 6; Page 43-47; 85bp; English.  
 CC The present invention provides protein and coding sequences from  
 CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and  
 CC able to replicate efficiently when transfected into a susceptible cell  
 CC line without reducing the growth rate of the cell line by more than 10  
 CC fold. The sequences are useful for screening for anti-HCV therapeutics,  
 CC for detecting antibodies to HCV in a biological sample such as blood,  
 CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,  
 CC for deriving authentic HCV components such as replication-competent non-  
 CC infectious, replication-defective infection-component, and replication-  
 CC defective non-infectious HCV, in gene therapy or gene vaccination  
 CC targeted to hepatic tissue for treating an animal infected or susceptible  
 CC to HCV infection and for studying HCV infection and propagation. The  
 CC present sequence is a clone of a fragment of the HCV genome which encodes  
 CC the core-neo and NS3 proteinase/helicase proteins  
 CC  
 SQ Sequence 7992 BP; 1648 A; 2369 C; 2243 G; 1732 T; 0 U; 0 Other;  
 Query Match 87.8%; Score 3291.4; DB 6; Length 7992;  
 Best Local Similarity 92.4%; Pred. No. 0;  
 Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 541 ACAGCCACCAAGCAGGAGCGCTGCTCCCGTGTGAGTCCAGTGGCGAGCCCTT 600  
 Db 3937 ACAGCCACCAAGCAGGAGCGCTGCTCCCGTGTGAGTCCAGTGGCGAGCCCTT 3996  
 QY 601 GAGACCTTCTGGGCGAAACATGTGGAATTCATCAGCGGGATACAGTACTTAAAGCAGC 660  
 Db 3997 GAGACCTTCTGGGCGAAACATGTGGAATTCATCAGCGGGATACAGTACTTAAAGCAGC 4056  
 QY 661 TTGTCCACTGCTGCTGGGAATCCGCGATTCATCATGAGCGTTCCAGAGCTCTGTC 720  
 Db 4057 TTGTCCACTGCTGCTGGGAATCCGCGATTCATCATGAGCGTTCCAGAGCTCTGTC 4116  
 QY 721 ACTAGCCCGCTCAACCAATATCACTCTGCTTAACTCTGCGGGAGTGAAGCC 780  
 Db 4117 ACCAGCCCGCTCAACCAATATCACTCTGCTTAACTCTGCGGGAGTGAAGCC 4176  
 QY 781 GCCCAACTGCTCTCCCAAGTGTGCTTCAAGCTTTGTAGAGCGCCGATTTGCTGTGC 840  
 Db 4177 GCCCAACTGCTCTCTCCCAAGTGTGCTTCAAGCTTTGTAGAGCGCCGATTTGCTGTGC 4236  
 QY 841 GCTGTTGGAGCATAGGCGCTTGGGAAGTGTGGAACATCTTGGCGGGCTATGAGCA 900  
 Db 4237 GCTGTTGGAGCATAGGCGCTTGGGAAGTGTGGAATTTTGGCAGGTTATGAGCA 4296  
 QY 901 GAGTGGCAGGCGCGCTGCTGCTTAAAGTTCATGAGCGCGAATGCTTCCACCGAG 960  
 Db 4297 GAGTGGCAGGCGCGCTGCTGCTTAAAGTTCATGAGCGCGAATGCTTCCACCGAG 4356  
 QY 961 GACTGTGTTAATCTTACTCTCCCTGCACTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1020  
 Db 4357 GACTGTGTTAATCTTACTCTCCCTGCACTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 4416  
 QY 1021 TGGGAGCGCATACTGCTGCGGACGCTGCTGCGAGGGAAGGGGCTTGCATGAGTAAC 1080  
 Db 4417 TGGGAGCGCATACTGCTGCGGACGCTGCTGCGAGGGAAGGGGCTTGCATGAGTAAC 4476  
 QY 1081 CGGCTATAGCGTTCCTCTGCGGGGTAACAGTTCCTCCCGACGACTATGCGCAGAG 1140  
 Db 4477 CGGCTATAGCGTTCCTCTGCGGGGTAACAGTTCCTCCCGACGACTATGCGCAGAG 4536  
 QY 1141 AGGAGCGCGGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
 Db 4537 AGGAGCGCGGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4596  
 QY 1201 AAGAGCTCCACGAGGATTAACAGAGACTGCTCCAGCCCTGCTCGGGCTGTGAGCTA 1260  
 Db 4597 AAGAGCTCCACGAGGATTAACAGAGACTGCTCCAGCCCTGCTCGGGCTGTGAGCTA 4656  
 QY 1261 AGGAGTGTGGGACTGATGATGACAGTTTGGCTGACTGAACCTGGCTCCAGTCC 1320  
 Db 4657 AGGAGTGTGGGACTGATGATGACAGTTTGGCTGACTGAACCTGGCTCCAGTCC 4716  
 QY 1321 AAGCTCTGCTCGGATTAACCGGAGTCCCTTTTCTCATGCAACGTTGGGTAACAGGGG 1380  
 Db 4717 AAGCTCTGCTCGGATTAACCGGAGTCCCTTTTCTCATGCAACGTTGGGTAACAGGGG 4776  
 QY 1381 GTCTGCGGGAGACGGCATCATGACACCACTGCTCATGTGAGACACAGTCAACCGGA 1440  
 Db 4777 GTCTGCGGGAGACGGCATCATGACACCACTGCTCATGTGAGACACAGTCAACCGGA 4836  
 QY 1441 CATGTCAAAAAGGTTCCATGAGGATCGTTGGGCTTAACCTGTATGTAACATGTGGCAT 1500  
 Db 4837 CATGTCAAAAAGGTTCCATGAGGATCGTTGGGCTTAACCTGTATGTAACATGTGGCAT 4896  
 QY 1501 GGAATCTTCCCATCAACGATACCAAGGCGCCCTGACGCGCTCCCAAGCGCCCAAC 1560  
 Db 4897 GGAATCTTCCCATCAACGATACCAAGGCGCCCTGACGCGCTCCCAAGCGCCCAAC 4956  
 QY 1561 TATTCAGGGCGCTGTGCGGGTGTGCTGTAAGAGTATGTAAGGTTACGCGGGTGGG 1620  
 Db 4957 TATTCAGGGCGCTGTGCGGGTGTGCTGTAAGAGTATGTAAGGTTACGCGGGTGGG 5016  
 QY 1621 GATTTCCACTAGTACGAGCATGACCACTGACACATGTAATAATGCCCGGTCCAGTTCCA 1680

||||| 5017 GATTTCACACTGACGCGCATGACACATGCAAGTAAAGTCCCGGTGTGAGGTTCCG 5076  
||||| 1681 GCGCCGGAATTTCTTCA CAGAAATGATGGGGTGCAGGTCGACAGTTACGCTCCGGCGTCG 1740  
5077 GCGCCGGAATTTCTTCA CAGAAATGATGGGGTGCAGGTCGACAGTTACGCTCCGGCGTCG 5136  
1741 AAACCTCTCCTACGAGGAGAGTCAATTCAGAGTGGGCTCAACCAATATCCTGTGTGG 1800  
5137 AAACCTCTCCTACGAGGAGAGTCAATTCAGAGTGGGCTCAATCAATCAATCCTGTGTGG 5196  
1801 TCGAGTCTCCATGCGAGGCCGAA CCGGATGTAGAGTGTCACTTCATGCTCACCGAC 1860  
5197 TCACAGCTCCATGCGAGGCCGAA CCGGATGTAGAGTGTCACTTCATGCTCACCGAC 5256  
1861 CCTCCCAATCA CAGAGAGAGGCTAAGCGAGGTGGCCAGGGGGTCTCCCGCTCC 1920  
5257 CCTCCCAATCA CAGAGAGAGGCTAAGCGAGGTGGCCAGGGGGTCTCCCGCTCC 5316  
1921 TTGGCAGCTCTT CAGCTAGCCAGTTGTCTGCGCTTCTCTGAAGGCGACATATAC 1980  
5317 TTGGCAGCTCTT CAGCTAGCCAGTTGTCTGCGCTTCTCTGAAGGCGACATATAC 5376  
1981 CAAATGACTTCCCAAGCGCTGACCTCATGAGGCCAACTCTGTGGCGGCAATGAGATG 2040  
5377 CGTATGACTTCCCGGAGCGCTGACCTCATGAGGCCAACTCTGTGGCGGCAAGATG 5436  
2041 GGGGGGGAATTCACCGCGGTGAGTCA GAGAACAAGTATTAATCCGTGACTCTTTCGAC 2100  
5437 GGGGGGAATTCACCGCGGTGAGTCA GAGAACAAGTATTAATTTGGACTCTTTCGAG 5496  
2101 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTGTCCCTCCGGCGAGATCTTGGCGAAA 2160  
5497 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTGTCCCTCCGGCGAGATCTTGGCGAAG 5556  
2161 TCCAAAGAAATTTCCCA CCAAGGATGCGCGCATGGGCAAGCCGGATTACAACCTTCGCTG 2220  
5557 TCCAGGAATTTCCCTCGAGCGATGCGCATATGGGCAAGCCGGATTACAACCTTCGCTG 5616  
2221 CTGAGATCTTGGAAGGCGCGGACTAGTCCCTTCAATGGTATCATGGGTGCCCATGCGCA 2280  
5617 TTTAGATCTTGGAAGGCGCGGACTAGTCCCTTCAATGGTATCATGGGTGCCCATGCGCG 5676  
2281 CCTACTAAGA CCGCTCTATATCAACTCA CCGAGAGAGAGACAGTTGTTCTGACAGAA 2340  
5677 CTTGCGAAGGCGCTTCCGATACACTTCA CCGAGAGAGAGAGAGGTTGTTCTGACAGAA 5736  
2341 TCCACCGTGTCTTGTGCTGCGGAGCTTGC CAAAGGCTTTCGGTAGCTTCGAAACG 2400  
5737 TTTACCGTGTCTTGTGCTTGGCGGAGCTCG CCAAAAGACCTTCGGCAGCTCCGAAATCG 5796  
2401 TCGGCGGTGCA CAGCGGCAAGGCAAGCCCTCTCTTA CCAACCTTCCGACAGCGGGA 2460  
5797 TCGGCGGTGCA CAGCGGCAAGGCAAGCCCTCTCTTA CCAACCTTCCGACAGCGGGA 5856  
2461 GCGAGATCTAGCGTTAGTGTATTTCTCTCAATG CCCCCCTTGAAGGGGAGCGCGGGAGAT 2520  
5857 GCGAGATCTAGCGTTAGTGTATCTCTCAATG CCCCCCTTGAAGGGGAGCGCGGGAGAT 5916  
2521 CCGCATCTAGCGAGCGGCTTTGGTCTTACCGTGA GAGAGCGCGGTGAAGAGTGTGTC 2580  
5917 CCGCATCTAGCGAGCGGCTTTGGTCTTACCGTGA GAGAGAGTGTGAAGAGTGTGTC 5976  
2581 TGTGCTGATGTCTTACACATGAGCAGGCGCTTGA TCAAGCCATCGCTGCGGAGAA 6036  
5977 TGTGCTGATGTCTTACACATGAGCAGGCGCTTGA TCAAGCCATCGCTGCGGAGAA 6036  
2641 AGGAAGTGCCTCAACAGCGTTGAGCAACTTTGCTG GGTGACCAACATGCTTAC 2700  
6037 ACCAAGTGCCTCAACATGCACTGAGCAACTTTGCTG GGTGACCAACATGCTTAC 6096  
2701 GCTACCAATCCGCGAGCGCAAGCGGAGAGAGTCA CTTTGA CAGACTGCAA 2760  
|||||

Db 6097 GCTACAACTCTCGACGCGCAAGCCTGCGGCAAGAAAGGTACCTTTGACAGACTGCAG 6156  
Qy 2761 ATCTGAGAGATCACTACCAAGACGTGCTCAAGAGATGAAAGCGAGCGTCCACAGTT 2820  
Db 6157 GTCTGAGAGACCACTACCGGAGCGTGTCAAGAGATGAAAGCGAGCGTCCACAGTT 6216  
Qy 2821 AAGGCTAGCTTCTATCAGTAAAGAGGCTGCAAGCTGACCGCCCCCACTTCCGCGCAA 2880  
Db 6217 AAGGCTAACTTCTATCCGTGAGAGAGGCTGTAAAGCTGACCGCCCCCACTTCCGCGCAA 6276  
Qy 2881 TCTAAATTTGGCTAAGGGGCAAGAGCGTCCGGAACCTATCCAGCAAGGCGATTAAACAC 2940  
Db 6277 TCTAAATTTGGCTAAGGGGCAAGAGCGTCCGGAACCTATCCAGCAAGGCGATTAAACAC 6336  
Qy 2941 ATCCGCTCCGTGAGGAGACTTGTGTGAAGA CACTGAACA CCAATTGACACCAACATC 3000  
Db 6337 ATCCGCTCCGTGAGGAGACTTGTGTGAAGA CACTGAAGA CCAATTGACACCAACATC 6396  
Qy 3001 ATGGCAAAAATGAGGTTTCTGCGTCCAAC CAGAGAGAGAGCGCGCAAGCCAGCTCGC 3060  
Db 6397 ATGGCAAAAATGAGGTTTCTGCGTCCAAC CAGAGAGAGAGCGCGCAAGCCAGCTCGC 6456  
Qy 3061 CTTATCGTGTCCAGACTTGGGGGTCGCTGTGAGAGTGGCGGAGAAAATGGCCCTATGACG 3120  
Db 6457 CTTATCGTGTCCAGACTTGGGGGTCGCTGTGAGAGTGGCGGAGAAAATGGCCCTATGAGT 6516  
Qy 3121 GTCTCAACCTCCCTCAGGCTGTGATGGAGCTCTCTGTACGATTCAGATTTCTCTGGA 3180  
Db 6517 GTCTCAACCTCCCTCAGGCGGTGATGGAGCTCTCTGTACATGATTCAGATTTCTCTGGA 6576  
Qy 3181 CAGCGGTCGAGTTCCTGTGTAA CCGCTGGAATCAAGAGACCCCTATGGGCTTTGCA 3240  
Db 6577 CAGCGGTCGAGTTCCTGTGTAA CCGCTGGAATCAAGAGACCCCTATGGGCTTTGCA 6636  
Qy 3241 TATGACACCCGCTGTTTATCAACAGTCA CTAAGAGAGATCATCCGTGTAGAGAGTCA 3300  
Db 6637 TATGACACCCGCTGTTTATCAACAGTCA CTAAGAGAGATCATCCGTGTAGAGAGTCA 6696  
Qy 3301 ATTATCAATGTTGTGACTTGGCCCCCGAAGCGAG CAGAGCTGAGTGTGCTCACAGAG 3360  
Db 6697 ATTATCAATGTTGTGACTTGGCCCCCGAAGCGAG CAGAGCTGAGTGTGCTCACAGAG 6756  
Qy 3361 CCGCTTTATATCGGGGTCCTCTGACTAAAT CCAAGGCGAGAACTGCGGCTATGCGCGG 3420  
Db 6757 CCGCTTTATATCGGGGTCCTCTGACTAAAT CCAAGGCGAGAACTGCGGCTATGCGCGG 6816  
Qy 3421 TGTGCGCGGAGCGGCGTGTGAGAGACTAGTGGGTAAATACCTTCACATGTTACTTGAAG 3480  
Db 6817 TGTGCGCGGAGCGGCGTGTGAGAGACTAGTGGGTAAATACCTTCACATGTTACTTGAAG 6876  
Qy 3481 GCGCTGAGCGCTGTGAGCGTGAAGCTCAAGGACTG CACAGATGCTGCTGTGCGGAGAG 3540  
Db 6877 GCGCTGAGCGCTGTGAGCGTGAAGCTGAGAGCTCAGAGACTG CACAGATGCTGCTGTGCGGAGAG 6936  
Qy 3541 GCGCTGTGCTTATCTGTAGAGCGCGGAA CCAAGAGAGACGCGGCGAGCTTACAGATGTC 3600  
Db 6937 GACCTTGTGCTTATCTGTAGAGCGCGGAA CCAAGAGAGACGCGGCGAGCTTACAGATGTC 6996  
Qy 3601 TTCAAGAGGCTATGACTAGAGTCTGCCCCCCCCCGGGGACCGCGCCCAACCAAGATAC 3660  
Db 6997 TTCAAGAGGCTATGACTAGAGTCTGCCCCCCCCCGGGGACCGCGCCCAACCAAGATAC 7056  
Qy 3661 GACCTGAGGTTGATATCATGATGCTCCCAATGTGTGCGGTGCGGACGATGATCTGGC 3720  
Db 7057 GACCTGAGGTTGATATCATGATGCTCCCAATGTGTGCGGTGCGGACGATGATCTGGC 7116  
Qy 3721 AAAAGGTATACTACTCAACCCGTGACCC 3749  
Db 7117 AAAAGGTATACTACTCAACCCGTGACCC 7145  
RESULT 8  
AAA98967



QY 1381 GTCTGCGGGGAGACGGCATCATGACACCCTGTCTCATGTGAGACAGATACCGGA 1440  
DB 4789 GTCTGCGGGGAGACGGCATCATGACACCCTGTCTCATGTGAGACAGATACCGGA 4848  
QY 1441 CATGTCAAAAACGGTTCATGAGAGATCTGGGCTTAAGACCTGTAGTAACATGTGGCAT 1500  
DB 4849 CATGTCAAAAACGGTTCATGAGAGATCTGTGGGCTTAAGACCTGTAGTAACATGTGGCAT 4908  
QY 1501 GGAACATTTCCCATCAACGATACACCAAGGGGCCCTGACGGCCCTCCCGAGGCCAAAC 1560  
DB 4909 GGAACATTTCCCATCAACGATACACCAAGGGGCCCTGACGGCCCTCCCGAGGCCAAAT 4968  
QY 1561 TATTCAGAGGCGGTGTGCGGGGTGTCTGAGAGATCGTGAAGTTACGCGGTGGG 1620  
DB 4969 TATTCAGAGGCGGTGTGCGGGGTGTCTGAGAGATCGTGAAGTTACGCGGTGGG 5028  
QY 1621 GATTTCCATAGTGTACAGAGCATGACCTGACACGTAATAATGCCGTGCCAGTTCCA 1680  
DB 5029 GATTTCCATAGTGTACAGAGCATGACCTGACACGTAATAATGCCGTGCCAGTTCCG 5088  
QY 1681 GCGCCCGAATTTCTCAAGAAATGGAAATGGGGTGGGTGACAGATACGCTCCGGGGTGC 1740  
DB 5089 GCGCCCGAATTTCTCAAGAAATGGAAATGGGGTGGGTGACAGATACGCTCCAGCGTGC 5148  
QY 1741 AAACCTCTCTCTACCGGAGAGAGTCAACATTCAGAGTCCGGCTCAACCAATACCTGTGGG 1800  
DB 5149 AAACCTCTCTCTACCGGAGAGAGTCAACATTCCTGTGGGGCTCAATCAATACCTGTGGG 5208  
QY 1801 TCGCAGCTCCCATGCGAGACCCGGAACCGGATGTACAGTGTCTCATCTTCATGTCACCGAC 1860  
DB 5209 TCGCAGCTCCCATGCGAGACCCGGAACCGGATGTACAGTGTCTCATCTTCATGTCACCGAC 5268  
QY 1861 CCTCCCAATCAACAGAGAGAGCGGCTAAGCGAGGTGGCGAGGGGGTCTCCCGCTCC 1920  
DB 5269 CCTCCCAATCAACAGAGAGAGCGGCTAAGCGAGGTGGCGAGGGGGTCTCCCGCTCC 5328  
QY 1921 TTGGCAGAGCTCTTCAAGTACACAGTTGTCTGCGCTTCTCGAAGCGACATACATTTACC 1980  
DB 5329 TTGGCAGAGCTCTTCAAGTACACAGTTGTCTGCGCTTCTCGAAGCGACATACATGACCTAC 5388  
QY 1981 CAATAATGATCTCCCAAGAGCTGTAACCTCATGAGGCCAATCTCTGTGGCGGACATGAGATG 2040  
DB 5389 CGTATATGATCTCCCGGAGCGCTGACCTCATGAGGCCAATCTCTGTGGCGGAGAGATG 5448  
QY 2041 GGGGGGGAAGTATCCCGCGTGAAGTCAAGAAAGATTAATCTTGAGCTCTTTCGAC 2100  
DB 5449 GGGGGGGAAGTATCCCGCGTGAAGTCAAGAAAGATTAATCTTGAGCTCTTTCGAG 5508  
QY 2101 CCGCTCCGAGCGGAGAGATGAGCGGAGATGTCCGTCCGGCGGAGATCTTGCGGAAA 2160  
DB 5509 CCGCTCCGAGCGGAGAGATGAGCGGAGATGTCCGTCCGGCGGAGATCTTGCGGAAA 5568  
QY 2161 TCGAAGAAATTTCCCAAGAGATGCCGACATGGGACAGCCCGGATTAACAATCTTCGCTG 2220  
DB 5569 TCGAAGAAATTTCCCTGAGGAGATGCCATATGGGACAGCCCGGATTAACAATCTTCACCTG 5628  
QY 2221 CTGGAATCTTGGAAGGCGCGGACCTAATGCTCCCTCAATGGTAATAGGTGCGCCACGCGCA 2280  
DB 5629 TTAAGATCTTGGAAGGCGCGGACCTAATGCTCCCTCAATGGTAATAGGTGCGCGCATTTGCGCG 5688  
QY 2281 CTTACTAAGACCCCTCTATACCACTCAACGAGAGAGAGACAGTTGTTTGAAGAA 2340  
DB 5689 CTTACTAAGACCCCTCTGATACCACTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5748  
QY 2341 TCGACCGTGTCTTGTGCGCTGCGGAGCTTGCAAAAGGCTTTCGATGCTCGAAGCG 2400  
DB 5749 TCTACCGTGTCTTGTGCGCTTGGCGAGCTGCGCAAAAGACCTTTCGCAAGCTCGAATCG 5808  
QY 2401 TCGGCGGTGAGAGCGGACAGCGGACCGCCCTTCTTACCAACCTCTCGACAGAGCGGGA 2460  
DB 5809 TCGGCGGTGAGAGCGGACAGCGGACCGCCCTTCTTACCAACCGCTCTCGACAGAGCGGAC 5868

QY 2461 GAGAGATCTGACGTTGATGTGTAATTCCTCATATGCCCCCTTGTAGGGGAGAGCGGGGAGAC 2520  
DB 5869 GCGGATCCGAGATTGAGTGTATCTCTCATATGCCCCCTTGTAGGGGAGAGCGGGGAGAT 5928  
QY 2521 CCGGATCTGAGGAGCGGGTCTTGTGTACCGTGAAGAGAGGCGGTGAGAGCGTGCAT 2580  
DB 5929 CCGGATCTGAGGAGCGGGTCTTGTGTACCGTGAAGAGAGGCTTATGATGAGAGCGTGCAT 5988  
QY 2581 TGTGTCTGATGTCTTACACATGAGACAGGCGCTGTGATACCGCATATGCCGTGCGAGGAA 2640  
DB 5989 TGTGTCTGATGTCTTACACATGAGACAGGCGCGCTGTATACCGCATATGCCGTGCGAGGAA 6048  
QY 2641 AGCAAGCTCCCATCAACCGGTTAGCAACTCTTGTCTGCTGTCACCAACAATGTCTTAC 2700  
DB 6049 ACCAAGCTCCCATCAATGACATGAGCAACTCTTGTCTGCTGTCACCAACAATGTGTCTAT 6108  
QY 2701 GCTTACCAATCCCGAGCGGCAAGCGGCGGAGAGAGAGTCACTTTGACAGACTGCAAG 2760  
DB 6109 GCTTACCAATCTTGCAGCGGCAAGCGCTGCGGAGAGAGAGTCACTTTGACAGACTGCAAG 6168  
QY 2761 ATCTGAGAGATCACTACAGGAGCGTGTCAAGAGATGAAAGCGAGCGTCCACAGTT 2820  
DB 6169 GTCTGAGAGACACACTACCGGAGCGTGTCAAGAGATGAAAGCGAGCGTCCACAGTT 6228  
QY 2821 AAGGCTAAGCTTCTATCAATGAGAGAGCGTGTCAAGCTGACCGCCCATATTCGCGCAAA 2880  
DB 6229 AAGGCTAAGCTTCTATCCGTGAGAGAGCGTGTCAAGCTGACCGCCCATATTCGCGCAAA 6288  
QY 2881 TCTAAATTTGGCTATGAGGCGCAAGAGCGTCCGGAACCTTACAGCAAGGCGATTAACAC 2940  
DB 6289 TCTAAATTTGGCTATGAGGCGCAAGAGCGTCCGGAACCTTACAGCAAGGCGGTTAACAC 6348  
QY 2941 ATCCGCTCCGTGTGAGAGACTTGTGAAAGCACTGAAACCAATTTGACACCAACATC 3000  
DB 6349 ATCCGCTCCGTGTGAGAGACTTGTGAAAGCACTGAAAGCACTGAAAGCACTGAAAGCACT 6408  
QY 3001 ATGGCAAAAATGAGGTTTTTCTGCTCAACAGAGAGAGAGGCGCAAGCCAGCTCCG 3060  
DB 6409 ATGGCAAAAATGAGGTTTTTCTGCTCAACAGAGAGAGAGGCGCAAGCCAGCTCCG 6468  
QY 3061 CTATATGTGTCTCCCAAGCTTTGGGGGTCGCTGTGTGAGAAAATGGCCCTTATGACGATG 3120  
DB 6469 CTATATGTGTCTCCCAAGCTTTGGGGGTCGCTGTGTGAGAAAATGGCCCTTATGACATGTG 6528  
QY 3121 GTCTCAACCTCCCTCAAGCTGTGATGGGCTCTCTGATGAGATTCAGATTTCTCTGGA 3180  
DB 6529 GTCTCAACCTCCCTCAAGCTGTGATGGGCTCTCTGATGAGATTCAGATTTCTCTGGA 6588  
QY 3181 CAGCGGGTGAAGTCTGTGTAACGCTGGAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
DB 6589 CAGCGGGTGAAGTCTGTGTAATGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6648  
QY 3241 TATGACACCGCGCTTTTGAATCAAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 3300  
DB 6649 TATGACACCGCGCTTTTGAATCAAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 6708  
QY 3301 ATTTATCAATGTGTGATCTTGGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
DB 6709 ATTTATCAATGTGTGATCTTGGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6768  
QY 3361 CCGCTTATATCGGGGGTCCCTGACTAAATTCAAAAGGCGAGAACTGCGGCTATTCGCGG 3420  
DB 6769 CCGCTTATATCGGGGGTCCCTGACTAAATTCAAAAGGCGAGAACTGCGGCTATTCGCGG 6828  
QY 3421 TGCCTGCGAGCGGCGTGTCTGACAGTATGCTGCGTAAATACCTTCATATTTACTTGAAG 3480  
DB 6829 TGCCTGCGAGCGGCGTGTCTGACAGTATGCTGCGTAAATACCTTCATATTTACTTGAAG 6888  
QY 3481 GCGCTGAGCGCTGTGAGAGTGAAGCTCCAGAGCTGACAGATGCTGCTGAGAGAGAG 3540  
DB 6889 GCGCTGAGCGCTGTGAGAGTGAAGCTTCAGAGCTGACAGATGCTGATATCGAGAGAG 6948  
QY 3541 GCGCTGTGTGATCTGTGAGAGCGGAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600



|||||  
Db 5125 CGGCTATAGCGTTCCTTCGCGGGGTAACAGTCTCCCGCAAGCATATGAGCTTGAG 5184  
QY 1141 AGGAGCGCGAGACAGTGTCACTCAGATCTCTCGAAGCTTACTATACCCAACTGGT 1200  
Db 5185 AGGAGCGTGCACAGTGTCACTCAGATCTCTCTAGTTTACATCACTCACTGCTG 5244  
QY 1201 AAGAGCTCCACAGTGAATTAAAGAGACTGTCTCAAGCCCTGCTCGGCTGTGGCTA 1260  
Db 5245 AAGAGCTTCACAGTGAATTAAAGAGACTGTCTCAAGCCCTGCTCGGCTGTGGCTA 5304  
QY 1261 AAGAGTGTGGAGCTGAGATATGACAGTTTGGCTGACTTAAAGACTTGGCTCGCTC 1320  
Db 5305 AAGAGTGTGGAGTGAATGATCAAGTGTGACTGATTTCAAGACTTGGCTCGCTC 5364  
QY 1321 AAGCTCTGCGGAGTATACCGGAGTCCCTTTTTCATGSCAAAGTGGGTCAAGGGG 1380  
Db 5365 AAGCTCTGCGGAGTATGCGGAGTCCCTTCTTCATGTCAAGTGGGTCAAGGGG 5424  
QY 1381 GTCTGCGGGAGACGAGCATCATGACAGACCACTGCTCATGTGAGACAGATCACCGG 1440  
Db 5425 GTCTGCGGGAGACGAGCATCATGACAGACCACTGCTCATGTGAGACAGATCACCGG 5484  
QY 1441 CATGTCAAAAACGTTTCATGAGATCTGTTGGCTTAAGACTGTAGTAACTATGGCAT 1500  
Db 5485 CATGTCAAAAACGTTTCATGAGATCTGTTGGCTTAAGACTGTAGTAACTATGGCAT 5544  
QY 1501 GGAACATTCCTCCATCAACGATACACAGGGGCGGCTGACGAGCCCTCCCGGAGCCAAAC 1560  
Db 5545 GGAACATTCCTCCATCAACGATACACAGGGGCGGCTGACGAGCCCTCCCGGAGCCAAAC 5604  
QY 1561 TATTCAGAGCGCTGTGTGCGGGTGTGCTGAGAGATGCTGAGAGTTTACGCGGATGGG 1620  
Db 5605 TATTCAGAGCGCTGTGTGCGGGTGTGCTGAGAGATGCTGAGAGTTTACGCGGATGGG 5664  
QY 1621 GATTTTCACTACGTGACAGACATGACATGACAGTAAATGCCCCGCTCCAGTTTCCA 1680  
Db 5665 GATTTTCACTACGTGACAGACATGACATGACAGTAAATGCCCCGCTCCAGTTTCCG 5724  
QY 1681 GCCCCCGAATTTCTTCAAGAAATGAGTGGGGTGGCTGACAGGATCGCTCCGGGCTGC 1740  
Db 5725 GCCCCCGAATTTCTTCAAGAAATGAGTGGGGTGGCTGACAGGATCGCTCCAGGCTGC 5784  
QY 1741 AAACCTCTCTTACGAGAGAGTCACTTCCAGGTGGGCTCAACCAATACCTGTGGG 1800  
Db 5785 AAACCTCTCTTACGAGAGAGTCACTTCTGGTGGGCTCAATCAATACCTGTGGG 5844  
QY 1801 TCGAGCTCTCCATGCGAGCCGAAACGAGATGACAGTCTCACTTCATGCTCACCGAC 1860  
Db 5845 TCGAGCTCTCCATGCGAGCCGAAACGAGATGACAGTCTCACTTCATGCTCACCGAC 5904  
QY 1861 CCGTCCCACTCAAGAGAGACGAGTAAAGGAGCTGAGCAAGGGGCTTCCCGCTCC 1920  
Db 5905 CCGTCCCACTCAAGGAGAGAGCTTAAAGGAGCTGAGCAAGGGGATTCCTCCCGCTCC 5964  
QY 1921 TTGGCAGCTCTTCACTAGTACCACTGTCTGCGCTTCTCGAAGGCACTATTAACC 1980  
Db 5965 TTGGCAGCTCTTCACTAGTACCACTGTCTGCGCTTCTCGAAGGCACTATTAACC 6024  
QY 1981 CAAAATGACTTCCAGACGCTGACCTCATGAGGCAACTCTGTGGCGGCACTAGAGT 2040  
Db 6025 CAGTATGACTTCCCGGAGCTGACCTCATGAGGCAACTCTGTGGCGGAGAGT 6084  
QY 2041 GGGGGGAGCTTCCCGGAGCTGACAGAAAGTAAATCCGAGACTCTTTCGAC 2100  
Db 6085 GGGGGGAGCTTCCCGGAGCTGACAGAAAGTAAATCCGAGACTCTTTCGAG 6144  
QY 2101 CCGCTCGAGCGAGAGATGAGCGGAGAGTGTCCCTCCGGCGAGATCTGTGGGAAA 2160  
Db 6145 CCGCTCGAGCGAGAGATGAGAGGAGTATCCGTTCCGGGAGATCTGTGGGAGG 6204  
QY 2161 TCGAAGAAATTCACAGGAGATGCCGATGGGCAAGCCCGGATTAACAACCTTCGCTG 2220  
|||||

Db 6205 TCGAAGAAATTCCTCGAGCATGCCCATATGAGGACGCCCGGATTAACAACCTTCACATG 6264  
QY 2221 CTGAGATCTCGAAGAGCCCGGAGTACGTCCCTCCAGTGTATCATGGGTGCCACTGCCA 2280  
Db 6265 TTAGATCTCGAAGAGCCCGGAGTACGTCCCTCCAGTGTATCATGGGTGCCACTGCCG 6324  
QY 2281 CCTACTAGACCCCTCTCTATATCACTCCACGAGAGAGAGACAGTTGTTGACAGAA 2340  
Db 6325 CCGCAAGAGCCCTCTCGATACACTCCACGAGAGAGAGACAGTTGTTGCTGTACAGAA 6384  
QY 2341 TCGACCGTGTCTTCCGCTGCGGAGCTTGCACAAAGCTTTGGGTAGCTCCGAACCG 2400  
Db 6385 TCTACCGTGTCTTCCGCTTGGGAGCTTGCACAAAGACTTGGGCACTTCCGATG 6444  
QY 2401 TCGGCGGTGACAGAGGACAGGCAACCGGCTCTCTGACCAACCTTCGACAGACGCGG 2460  
Db 6445 TCGGCGGTGACAGAGGACAGGCAACCGGCTCTCTGACCAACCGCTCGACAGCGGAC 6504  
QY 2461 GAGAGATCGACGTTGATGCTATTCCTCATGCCCCCTTTGAGGGGAGCCGGGGAC 2520  
Db 6505 GCGGATCCGAGTGAAGTGTATCTCTCATGCCCCCTTGAGGGGAGCGCGGGAT 6564  
QY 2521 CCGGATCGAGAGCGGCTTGTGTACCGTGAAGTGAAGGCGGATGAGACGTCGTC 2580  
Db 6565 CCGGATCGAGAGCGGCTTGTGTACCGTGAAGTGAAGGCGGATGAGACGTCGTC 6624  
QY 2581 TGTGCTCGATGTCTTCAATGACAGAGCGCTGTGATCAGCGCATGCGCTCGAGAGAA 2640  
Db 6625 TGTGCTCGATGTCTTCAATGACAGAGCGCGCTGTGATCAGCGCATGCGCTCGAGAGAA 6684  
QY 2641 AGCAAGCTGCCATCAACGCGTTGAGCAACTCTTGTGCTGCTACCAACAATGTCTAC 2700  
Db 6685 ACCAAGCTGCCATCAATCAATGAGCAACTCTTGTGCTGCTGCTGCTGCTGCTGCT 6744  
QY 2701 GCTACCAATCCCGAGCGGACAGCGAGGAGGAGAGAGTCACTTTGACAGACTGCA 2760  
Db 6745 GCTACCAATCTCGAGCGGAGCGGAGCGGAGGAGAGAGTCACTTTGACAGACTGCA 6804  
QY 2761 ATCTGAGAGATCACTACAGAGACGTGTCAAGAGATGAGGAGGAGGCTTCAAGATT 2820  
Db 6805 ATCTGAGAGATCACTACAGAGACGTGTCTAAGAGATGAGGAGGAGGCTTCAAGATT 6864  
QY 2821 AAGGCTAAGCTTCTATCAATGAGAGAGCTGTGCAAGCTGACGCCCCCACTTGGCTCAA 2880  
Db 6865 AAGGCTAAGCTTCTATCCGTGAGAGAGCTGTGCAAGCTGACGCCCCCACTTGGCTCAA 6924  
QY 2881 TCTAAATTTGGCTATGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940  
Db 6925 TCTAAATTTGGCTATGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6984  
QY 2941 ATCCGCTCCGTGTGAGAGACTTGTGGAAGACATGAACAACAATTGACACCAATC 3000  
Db 6985 ATCCGCTCCGTGTGAGAGACTTGTGGAAGACATGAACAACAATTGACACCAATC 7044  
QY 3001 ATGGCAAAAATGAGGTTTTTCTGCTGCAACAGAGAGAGAGGCTCGCAAGCAAGTCCG 3060  
Db 7045 ATGGCAAAAATGAGGTTTTTCTGCTGCAACAGAGAGAGAGGCTCGCAAGCAAGTCCG 7104  
QY 3061 CTTATGCTGTCCCAACTTTGGGGGCTCGTGTGTGGAAGAAATGGCCCTATGAGCGT 3120  
Db 7105 CTTATGCTGTCCCAACTTTGGGGGCTCGTGTGTGGAAGAAATGGCCCTATGAGTGT 7164  
QY 3121 GTCTCAACCTCCCTCAGGCTGATGAGGCTCCTCGTACGATTCAGATTCCTCTGGA 3180  
Db 7165 GTCTCAACCTCCCTCAGGCTGATGAGGCTCCTCGTACGATTCAGATTCCTCTGGA 7224  
QY 3181 CAGCGGCTGAGTTCTGTGTGAACGCTTGAAATCAAAGAGACCCCTATGAGCTTTGCA 3240  
Db 7225 CAGCGGCTGAGATTCGTGTGAATGCTTGAAAGCAAGAAATGCCCTATGAGCTTTGCA 7284  
QY 3241 TATGACACCCGCTGTTTGAATCAACATGACCTGAGAAATGCACTGATGAGAGGTCA 3300  
Db 7285 TATGACACCCGCTGTTTGAATCAACATGACCTGAGAAATGCACTGATGAGAGGTCA 7344  
|||||



OY	3301	ATTATTCATGTTGTGACTTGGACCCCGGAGACAGACAGGGCATPAAGTCGCTACAAG	3360
Db	7345	ATCTACCAATGTTGTGACTTGGCCCCCGAAGCCACAGGCCCTAAGTCGCTACAAG	7404
OY	3361	CGGCTTTATTCGGGGGTCCTCCCTGACTAATTCAAAAGGGCAGAACTGCGGCTATCGCGG	3420
Db	7405	CGGCTTTATTCATCGGGGGGCCCCCTCGACTAATTTCTAAAGGGCAGAACTGCGGCTATCGCGG	7464
OY	3421	TGCGCGCGAGCGGGGTGCTGACCATAGCTGCGGCTATATACCCTCACTGTACTTTAG	3480
Db	7465	TGCGCGCGAGCGGGGTGACTGACACGACCGACTGCGGCTAATATACCCTCACTGTACTTTAG	7524
OY	3481	GCCTCTGACGCGCTGTGCGACTGCCAAGCTCCAGACCTGCAGATGCTGTGTGGGAGAC	3540
Db	7525	GCGCGTGGCGCTGTGCGAGCTGCCAAGCTCCAGACCTGCAGATGCTGTGTGGGAGAC	7584
OY	3541	GGCCTTTCGTTTATCTGTGAGAGCGCGGGAAACCCAGAGAGACGGCGGACGCTTACGACTC	3600
Db	7585	GACCTTTCGTTTATCTGTGAAAGCGCGGGAGCCCAAGAGAGACGAGCGGACGCTTACGCGCC	7644
OY	3601	TTCAACGAGGCTATGACTAGGTACTCTGCCCCCGGAGACCGCGCCCAACCAAGATAC	3660
Db	7645	TTCAACGAGGCTATGACTAGATCTCTGCCCCCGTGGGAGACCGGCCAATCCAGATATAC	7704
OY	3661	GACCTGAGTTGATPAACATCATGCTCTTCCATGTGTGAGTCGCGGCACGATGACTTGGC	3720
Db	7705	GACCTGAGTTGATPAACATCATGCTCTTCCATGTGTGAGTCGCGGCACGATGACTTGGC	7764
OY	3721	AAAAGGATACTACTCCACCGCGAGCC	3749
Db	7765	AAAAGGATACTACTCAACCGCGAGCC	7793

```

RESULT 10
ABK88595
ID      ABK88595 standard; DNA; 8638 BP.
XX
XX      ABK88595;
AC
XX      21-OCT-2002 (first entry)
DT
XX      Hepatitis C virus AFGK12-5'A replicon.
DE
XX      Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
KW      cell culture replication; ds.
XX      Hepatitis C virus.
OS
XX
XX      Key
FH      Location/Qualifiers
FH      replace(1,G)
FT      mutation
FT      /tag= a
FT      1802..8407
FT      /tag= b
FT      /product= "HCV NS2-5B"
FT      /note= "viral enzymes"
XX
XX      WO200252015-A2.
XX
XX      04-JUL-2002.
XX
XX      20-DEC-2001; 2001MO-CA001843.
XX
XX      22-DEC-2000; 2000US-0257857P.
XX      (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.
XX      Kukolj G, Pause A;
XX      PI
XX      WPI: 2002-575382/61.
XX      DR      P-PSDB; ABG30587.
XX      TX
XX      New self-replicating RNA molecules from Hepatitis C virus (HCV), which

```

Query Match	Best Local Similarity	87.8%	Score 3291.4	DB 6	Length 8638
Matches 3463	Conservative	0	Mismatches 286	Indels	Gaps 0
1	TGGAGAGGCGCTTTCACAGGCTTCAACCCAGCTGATGCGCACTTCTGTCGCCAACAAAG	60			
4046	TGGAGAGGCGCTTTCACAGGCTTCAACCCAGCTGATGCGCACTTCTGTCGCCAACAAAG	4105			
61	CAGGCAAGAGAACAACTTCCCTTACTGCTGTCGCGTACCAAGGCTACTGTGTGCGTATGGGCT	120			
4106	CAGGCAAGAGAACAACTTCCCTTACTGCTGTCGCGTACCAAGGCTACTGTGTGCGTATGGGCT	4165			
121	CAGGCGCCACCTCCATCATCATGGAGATCAATGAGAGTCTCATACGCGCTAAAGGCTACT	180			
4166	CAGGCTCCACTCCATCTGTGGAGCCAAATGAGAGTCTCATACGCGCTAAAGGCTACTAG	4222			
181	CTGGCGGGGCGCAACCTTCTGCTATATGAGCTGGGAGCGCTTCAAAAGAGGTCACCTTC	240			
4226	CTGGCGGGGCGCAACCTTCTGCTATATGAGCTGGGAGCGCTTCAAAAGAGGTCATCTACC	4285			
241	ACACACCCCATTAACCAAAATTCATCATGCGCATGCTATGACCGCATCTGAGAGTCTGTACG	300			
4286	ACACACCCCATTAACCAAAATTCATCATGCGCATGCTATGACCGCATCTGAGAGTCTGTACG	4345			
301	AGCACCTGGGCGCTGCTGGGCGGGGCTCTTGCAGCTCTGCGCTGGGTAATGCTTGAACA	360			
4346	AGCACCTGGGCGCTGCTGGGCGGGGCTCTTGCAGCTCTGCGCGGTATGCTTGAACA	4405			
361	GGCAGCGTGCATCTTGTGGTAGAGATCATCTTGTCCGGGCGCGGCTATTTGTTCCGAC	420			
4406	GGCAGCGTGCATCTTGTGGTAGAGATCATCTTGTCCGGGCGCGGCTATTTGTTCCGAC	4465			
421	AGGGAAGTCTCTAACAGAGAGTTCATGATGATGAGAAAGATGCGGCTGCACCTTCCCTTAC	480			
4466	AGGGAAGTCTCTAACAGAGAGTTCATGATGATGAGAAAGATGCGGCTGCACCTTCCCTTAC	4525			
481	ATCGACGAGGAAATGCAAGCTCCCGGACAGTTCAAGCAAAAAGCGCTCCGGTTGCTGGAG	540			
4526	ATCGACGAGGAAATGCAAGCTCCCGGACAGTTCAAGCAAAAAGCGCTCCGGTTGCTGGAA	4585			
541	ACAGCCACCAAGCAAGCGAGGCGCTGCTCCGCTGAGAGTCAAGATGCGAGCCCTT	600			
4586	ACAGCCACCAAGCAAGCGAGGCGCTGCTCCGCTGAGAGTCAAGATGCGAGCCCTT	4645			
601	GAGACCTTCTGGGCGAAACAATGTGAAATTCATACGCGGAGTACAGTACTTACAGAGC	660			
4646	GAGACCTTCTGGGCGAAACAATGTGAAATTCATACGCGGAGTACAGTACTTACAGAGC	4705			
661	TTGTTCACCTGCTGGGAAATCCGCGAATTCATCATGATGAGGCTTCAAGCCTCTGTC	720			

Db 4706 TTGTCACTCTGCTGGCAACCCCGGATGATCATCTGATGGCAATTACAGACCTCTATC 4765  
QY 721 ACTAGCCCGCTGACCAACCAATCAACCTCTGCTTAAACATCTGGGGGATGAGTAGCC 780  
Db 4766 ACCAGCCCGCTGACCAACCAATCAACCTCTGCTTAAACATCTGGGGGATGAGTAGCC 4825  
QY 781 GCCCACTGCTCCCCCAAGTCTGCTTACGCTTTGCTAGGCGCCGGCAATTGCTGTGCG 840  
Db 4826 GCCCACTGCTCTCCCAAGCGCTGCTTGTGCTTGTAGCGCCGGCAATCGCTGAGCG 4885  
QY 841 GCTGTGGGAGCATAGGCTTTGGGAAGTCTTGTGACATCTTTGGGGCTTATGAGCA 900  
Db 4886 GCTGTGGGAGCATAGGCTTTGGGAAGTCTTGTGATATTTTGGAGGATATGAGCA 4945  
QY 901 GAGTGGCAGCGCGCTCGTGCCCTTAAAGTATGAGCGGCAAAATGCTCCACCGAG 960  
Db 4946 GGGGTGGCAGCGCGCTCGTGCCCTTAAAGTATGAGCGGCAAGATGCTCCACCGAG 5005  
QY 961 GACCTGGTTAACTTAATCTCCCTGCAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1020  
Db 5006 GACCTGGTTAACTTAATCTCCCTGCAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 5065  
QY 1021 TGGGAGCCATTAATCTGCTGCGGACATGAGGCTCAAGGGAGGGGCTGTGAGTGAATGAAC 1080  
Db 5066 TGGGAGCCATTAATCTGCTGCGGACATGAGGCTCAAGGGAGGGGCTGTGAGTGAATGAAC 5125  
QY 1081 CGGCTATAGCGTTGCGCTCGCGGGGTTAACATGTTTCCCGCAGCACTATGTCAGAG 1140  
Db 5126 CGGCTATAGCGTTGCGCTCGCGGGGTTAACATGTTTCCCGCAGCACTATGTCAGAG 5185  
QY 1141 AGGAGCGCGCAGCAGTGTCACTCAGATCCTTCCGACCTTATCAATCACTGTTG 1200  
Db 5186 AGGAGCGCTGACAGTGTCACTCAGATCCTTCTAGTCTTACATCACTGAGCTGCTG 5245  
QY 1201 AAGAGGCTCAACCAATGAGATTAAGAGAGTGTCTCCAGCCCTGCTGCTGCTGCTGCTG 1260  
Db 5246 AAGAGGCTCAACCAATGAGATTAAGAGAGTGTCTCCAGCCCTGCTGCTGCTGCTGCTG 5305  
QY 1261 AGGAGTGTGGGACTGATATGCAAGTGTGGGCTGACTTAAAGCACTGAGCTCAAGTTC 1320  
Db 5306 AGAGATGTTTGGGATTTGGAATGCAAGTGTGGGCTGACTTAAAGCACTGAGCTCAAGTTC 5365  
QY 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCAACGCTGGGTTAAAGAGG 1380  
Db 5366 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCAACGCTGGGTTAAAGAGG 5425  
QY 1381 GTCTGGCGGGAGACCGGATCATGAGACCACTGCTCATGTGAGGACAGATCACCGGA 1440  
Db 5426 GTCTGGCGGGAGACCGGATCATGAGACCACTGCTCATGTGAGGACAGATCACCGGA 5485  
QY 1441 CATGTCAAAAACGGTTCATGAGGATCGTTGGGCTTAAGACCTGTAGTAAATATGTGAT 1500  
Db 5486 CATGTGAAAAACGGTTCATGAGGATCGTTGGGCTTAAGACCTGTAGTAAATATGTGAT 5545  
QY 1501 GGAACATTTCCCATCAACGATACCAACGAGCCCTGCAAGCTCTCCCGACGCGCAAC 1560  
Db 5546 GGAACATTTCCCATCAACGATACCAACGAGCCCTGCAAGCTCTCCCGACGCGCAAC 5605  
QY 1561 TATTTCAAGGCGCTGTGGCGGGTGTGCTGAGAGTACGTGAGGTTAAGCGGGGTGGG 1620  
Db 5606 TATTTCAAGGCGCTGTGGCGGGTGTGCTGAGAGTACGTGAGGTTAAGCGGGGTGGG 5665  
QY 1621 GATTTCACTACGAGAGGATGACCATGACCAAGCTAAATGCGCTGACAGTTCGA 1680  
Db 5666 GATTTCACTACGAGAGGATGACCATGACCAAGCTAAATGCGCTGACAGTTCGA 5725  
QY 1681 GCCCCCAATTTCTTACAGAGTGAATGGGTGCGCTGCAAGSTACGCTCCGCGCTGC 1740  
Db 5726 GCCCCCAATTTCTTACAGAGTGAATGGGTGCGCTGCAAGSTACGCTCCGCGCTGC 5785  
QY 1741 AAACCTTCTCTACGGAGAGAGTCAATTCAGGTCGGGCTCAACCAATCTGTGTTGG 1800  
Db 5786 AAACCTTCTCTACGGAGAGAGTCAATTCCTGTGCGGCTCAATCAATCTGTGTTGG 5845

QY 1801 TCGAGCTTCCATGCGAGGCCCGAACCAGATGAGAGTGTCACTTCCATGCTCACCGAC 1860  
Db 5846 TCACAGCTCCCATGCGAGGCCCGAACCAGATGAGAGTGTCACTTCCATGCTCACCGAC 5905  
QY 1861 CCTTCCCAATTAACAGCAGAGACCGCTAAGCGGCAAGGCTGCGAGGGGCTCCCCCTTC 1920  
Db 5906 CCTTCCCAATTAACAGCAGAGACCGCTAAGCGGCAAGGCTGCGAGGGGATCTCCCCCTTC 5965  
QY 1921 TTGGCAGCTTTCAGCTAGCCAGTGTGTGCGCTTCTCCAGAGCGACATATATTAAC 1980  
Db 5966 TTGGCAGCTTTCAGCTAGCCAGTGTGTGCGCTTCTCCAGAGCGACATATATTAAC 6025  
QY 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGCGGCAATGAGATG 2040  
Db 6026 CCAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGCGGCAATGAGATG 6085  
QY 2041 GCGCGGGAATTAACCGCGCTGAGATGAGAAACAGTGTATCTGTGACTTTTCGAC 2100  
Db 6086 GCGCGGGAATTAACCGCGCTGAGATGAGAAATAGTATGTAATTTTGAATCTTTTCGAG 6145  
QY 2101 CGGCTCCAGCGGAGGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTTGCGGAAA 2160  
Db 6146 CGGCTCCAGCGGAGGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTTGCGGAG 6205  
QY 2161 TCCAGAAATTTCCACACAGCGATGCCCGCATGAGCACGCGCGAATTAACACCTTCGCTG 2220  
Db 6206 TCCAGAAATTTCCCTGACAGATGCCCATATGAGGACGCGCGGATTAACACCTTCGACTG 6265  
QY 2221 CTGAGATCTTGAAGAGCCCGGACTTCGCTTCCAGTGTATCATGAGTGTGCCACTGCGCA 2280  
Db 6266 TTAGAGTCTTGAAGAGCCCGGACTTCGCTTCCAGTGTATCATGAGTGTGCCACTGCGCG 6325  
QY 2281 CCTACTAAGACCCCTCTTAACCACTTCCAGGAGGAGAGACAGTGTTCGACAGAA 2340  
Db 6326 CTGACAGGCCCCCTTCCATACCACTTCCAGGAGGAGAGAGAGTGTTCGTCAGAA 6385  
QY 2341 TCCACCGTGTCTTTCGCTGCGGAGCTTGCACAAAGCTTTTGTGATCTCGAACCG 2400  
Db 6386 TCTACCGTGTCTTTCGCTGCGGAGCTTGCACAAAGCTTTTGTGATCTCGAACCG 6445  
QY 2401 TCGGCGGTGAGACGCGGACCGGCAACCGGCCCTCTGACCAACCTCTCCAGCGCGGCA 2460  
Db 6446 TCGGCGGTGAGACGCGGACCGGCAACCGGCCCTCTGACCAACCTCTCCAGCGCGGCA 6505  
QY 2461 GCAGGATCTGAGGTTGAGTGTATTCTCTCATGCCCCCTTGAAGGGAGAGCGGAGGAC 2520  
Db 6506 GCAGGATCTGAGGTTGAGTGTATTCTCTCATGCCCCCTTGAAGGGAGAGCGGAGGAT 6565  
QY 2521 CCGGATCTGAGGAGCGGATCTTGTCTACCGTGAAGTGAAGGCGCGTGAAGACGTCGT 2580  
Db 6566 CCGGATCTGAGGAGCGGATCTTGTGTCTACCGTGAAGGAGGAGGCTGTGAAGACGTCGT 6625  
QY 2581 TGTGCTCGAATGTCTTACATGAGACGCGCTCTGATACCGCATGCGCTCGGAGGAA 2640  
Db 6626 TGTGCTCGAATGTCTTACATGAGACGCGCTCTGATACCGCATGCGCTCGGAGGAA 6685  
QY 2641 AGCAAGCTGCCATCAACGCGTTGAGCAACTCTTGTCTGCTGACCAACAATGCTTAC 2700  
Db 6686 AGCAAGCTGCCATCAACGCGTTGAGCAACTCTTGTCTGCTGACCAACAATGCTTAC 6745  
QY 2701 GCTACCAATCCCGCAGCGCAAGCCAGCGGCAAGAAAGTCACTTTTGAACAGCTGCA 2760  
Db 6746 GCTACCAATCTCGAGCGGAGCGGCTCGCGGCAAGAAAGTCACTTTTGAACAGCTGCA 6805  
QY 2761 ATCTGAGCGATCACTTACAGAGCGTGTCTCAAGAGATGAAGCGCAAGGCTCTCAAGTT 2820  
Db 6806 ATCTGAGCGATCACTTACAGAGCGTGTCTCAAGAGATGAAGCGCAAGGCTCTCAAGTT 6865  
QY 2821 AAGGCTAAGCTTATCAAGTGAAGAGCTGCAAGCTGACCGCCCACTTGGGCAAA 2880  
Db 6866 AAGGCTAAGCTTATCTATCTGTGAGGAGGCTGTGAAGCTGACCGCCCACTTGGGCAAA 6925





Db 4347 AGCACTGGGTGCTGTAGCGGAGTCTTAGCAGCTTGCGCGGTAATGCTTGACAACA 4406  
Qy 361 GGCAGCGTGTCAATGTGGGTAGATCATCTTGCCGGCGCGGCTAATGTTCCGAC 4420  
Db 4407 GGCAGGTGTCAATGTGGGTAGATCATCTTGCCGGAAGCGGCGATCATTTCCGAC 4466  
Qy 4421 AGGAAGTCTTACCAAGAGTTGATGAGTGAAGAGTGGCGTGGCACTTCCTTAC 4480  
Db 4467 AGGAAGTCTTACCGGGAGTTGATGAGTGAAGAGTGGCGCTTACCACTTCCTTAC 4526  
Qy 481 ATGACAGGAGATGAGCTCGCGAGCAGTTGACGAAAGCGCTCGGTTGCTGAC 540  
Db 4527 ATGAAACAGGAATGAGCTCGCGAACAATTAACAGAAAGCAATCGGTTGCTGCA 4586  
Qy 541 ACAGCACCAAGCAAGCGAGCGCTGCTCCGTTGGAGTCAAGTGGCGAGCTT 600  
Db 4587 ACAGCACCAAGCAAGGAGGCTGCTCCGTTGGAGTCAAGTGGCGAGCTT 4646  
Qy 601 GAGACTTCTGGGCGAAACATGTGAACTTCATCAGCGGATACATACCTTAAGCAGC 660  
Db 4647 GAAGCTTCTGGGCGAAGCATATGTGAATTCATCAGCGGATACAAATTTAGCAGGC 4706  
Qy 661 TTGTCACTCTGCTGGGAAATCCCGGATTCATCTGATGGCTTCAAGCTCTGTC 720  
Db 4707 TTGTCACTCTGCTGGCAACCCCGGATTCATCTGATGGCTTCAAGCTCTTATC 4766  
Qy 721 ACTAGCCGCTTCAACCAACCATCTCTCTTAAACATCTGCGGGAGTGGTAAAGC 780  
Db 4767 ACAGGCCGCTTCAACCAACCATCTCTCTTAAACATCTGCGGGAGTGGTAAAGC 4826  
Qy 781 GCCCAACTGCTCCCCAGTGTCTTCACTTGTGAGGCGCGCATTTGCTGTGC 840  
Db 4827 GCCCAACTGCTCTCCAGCGCTGCTTGTGTAAGCGCGCATGCTGAGCGC 4886  
Qy 841 GCTTTGGAGCAATAGGCTTGGGAAGTGTGTGACATCTTGGCGGCTTTAGAGA 900  
Db 4887 GCTTTGGAGCAATAGGCTTGGGAAGTGTGTGATATTTTGGAGGTATAGAGA 4946  
Qy 901 GGAATGGACAGCGCGCTGTGACCTTTAAGTATGAGCGGCGAAATGCCCTCACCGAG 960  
Db 4947 GGGTGGCAGCGCGCTGTGACCTTTAAGTATGAGCGGCGAAATGCCCTCACCGAG 5006  
Qy 961 GACCTGGTTAATTAATCTCCCTGCAATCTCTCTCTGAGTCCCTGTGTGCGGAGTGTG 1020  
Db 5007 GACCTGGTTAATTAATCTCCCTGCAATCTCTCTCTGAGTCCCTGTGTGCGGAGTGTG 5066  
Qy 1021 TGGGACAGCAATCTGCTGGGCACTGGGTTCAGAGGAGGGGCTGTGCTGTGATGAC 1080  
Db 5067 TGGGACAGCAATCTGCTGGGCACTGGGTTCAGAGGAGGGGCTGTGCTGTGATGAC 5126  
Qy 1081 CGGCTATAGCGTTGCTGCGGAGGTAACATGTTTCCCGCACGACTATGTGCCAGAG 1140  
Db 5127 CGGCTATAGCGTTGCTGCGGAGGTAACATGTTTCCCGCACGACTATGTGCCAGAG 5186  
Qy 1141 AGGACGCGCGAGCAGTGTCACTCAAGATCTTCTCGACCTTAATCAACCAACTGTTG 1200  
Db 5187 AGGACGCGCGAGCAGTGTCACTCAAGATCTTCTCAATCACTCACTGAGTGTG 5246  
Qy 1201 AAAGAGTCCACAGATGAGATTAACAGAGATGCTCCAGCGCTCTCGGCTGTGGCTA 1260  
Db 5247 AAAGAGTCCACAGATGAGATTAACAGAGATGCTCCAGCGCTCTCGGCTGTGGCTA 5306  
Qy 1261 AGGAGTGTGGAGTGGATGATGACAGTTTGGCTGACTTAAGACCTGGCTCGAGTCC 1320  
Db 5307 AGGAGTGTGGAGTGGATGATGACAGTTTGGCTGACTTAAGACCTGGCTCGAGTCC 5366  
Qy 1321 AAGCTCTGCGGATTAACGGAGTCCCTTTTTCATGCAACGTTGGGTAACAAGG 1380  
Db 5367 AAGCTCTGCGGATTAACGGAGTCCCTTTTTCATGCAACGTTGGGTAACAAGG 5426  
Qy 1381 GTTGGCGGGGAGACGATCATGACAGACCACTGCTCATGTGAGAGCAAGATCACCGGA 1440  
Db 5427 GTTGGCGGGGAGACGATCATGACAGACCACTGCTCATGTGAGAGCAAGATCACCGGA 5486

Qy 1441 CATGTCAAAAACGTTTCATGAGGATGTTGGGCTTAAGACTGTATGTAATGTGGCAT 1500  
Db 5487 CATGTCAAAAACGTTTCATGAGGATGTTGGGCTTAAGACTGTATGTAATGTGGCAT 5546  
Qy 1501 GGAACATTTCCCATCAAGCATACACAGGCGCTTGCAGCGCTTCCAGCGCCAAAC 1560  
Db 5547 GGAACATTTCCCATCAAGCATACACAGGCGCTTGCAGCGCTTCCAGCGCCAAAC 5606  
Qy 1561 TATTCAGGAGCGTGTGGGAGTGGCTGTGAGAGTATGTAAGGTTATGCGCGGTGGG 1620  
Db 5607 TATTCAGGAGCGTGTGGGAGTGGCTGTGAGAGTATGTAAGGTTATGCGCGGTGGG 5666  
Qy 1621 GATTTCACTACGTGAAGAGATGACACTGACCAAGTAAATGCCGTGCAAGTTTCA 1680  
Db 5667 GATTTCACTACGTGAAGAGATGACACTGACCAAGTAAATGCCGTGCAAGTTTCCG 5726  
Qy 1681 GCCCGGAATTTCTTCAAGAGTGGATGGGCTGTGACAGATGAGCTCCGCGTGC 1740  
Db 5727 GCCCGGAATTTCTTCAAGAGTGGATGGGCTGTGACAGATGAGCTCCGCGTGC 5786  
Qy 1741 AAACCTCTCTACGGGAGAGGTCAATTCAGGTGGGCTCAACCAATACCTGTTGG 1800  
Db 5787 AAACCTCTCTACGGGAGAGGTCAATTCAGGTGGGCTCAACCAATACCTGTTGG 5846  
Qy 1801 TGCAGCTCCATGCGAGCGGAAACCGATGTAGAGTGTCACTTCAATGCTCACCGAC 1860  
Db 5847 TCACAGCTCCATGCGAGCGGAAACCGATGTAGAGTGTCACTTCAATGCTCACCGAC 5906  
Qy 1861 CCTCCCATCATCAGCAGAGATGCTTAAAGCGAGCTTGGCCAGGAGGTTCTCCCTTCC 1920  
Db 5907 CCTCCCATCATCAGCAGAGATGCTTAAAGCGAGCTTGGCCAGGAGTCTCCCTTCC 5966  
Qy 1921 TTGGCAGCTCTTACGCTAGCCAGTGTGTGCGCTTCTCTGAAAGCGCATACATTAC 1980  
Db 5967 TTGGCAGCTCTTACGCTAGCCAGTGTGTGCGCTTCTCTGAAAGCGCATACATTAC 6026  
Qy 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGCGGCAATGAGAT 2040  
Db 6027 CCAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGCGGCAATGAGAT 6086  
Qy 2041 GCGGGGGAATTAACCGGCTGTGAGTGAAGAACAGTGTATCTTGAATCTTTGAC 2100  
Db 6087 GCGGGGGAATTAACCGGCTGTGAGTGAAGAACAGTGTATCTTGAATCTTTGAC 6146  
Qy 2101 CGGCTCGAGGAGGAGATGAGGAGGTAAGTCCGTCGCGGAGATCTTGGCGGAA 2160  
Db 6147 CGGCTCGAGGAGGAGATGAGGAGGTAAGTCCGTCGCGGAGATCTTGGCGGAA 6206  
Qy 2161 TCCAAAGAAATTCACACAGCGATGCCCGATGGGCAAGCCGGAATTAACCTTCCGCTG 2220  
Db 6207 TCCAAAGAAATTCACACAGCGATGCCCGATGGGCAAGCCGGAATTAACCTTCCGCTG 6266  
Qy 2221 CTGAGTCTTGAAGGCGCGGACTATGCTTCTCAAGTGTATGAGTGGTCCGACTGCA 2280  
Db 6267 CTGAGTCTTGAAGGCGCGGACTATGCTTCTCAAGTGTATGAGTGGTCCGACTGCA 6326  
Qy 2281 CCACTAAGACCTCTCTAATCAACTCAAGGAGGAAAGGACAGTTGTTGACAGAA 2340  
Db 6327 CCACTAAGACCTCTCTAATCAACTCAAGGAGGAAAGGACAGTTGTTGACAGAA 6386  
Qy 2341 TCCACGCTGCTTCTGCGCTGCGAGTGTGCAAAAGCTTTCGATGCTCGAACC 2400  
Db 6387 TCCACGCTGCTTCTGCGCTGCGAGTGTGCAAAAGCTTTCGATGCTCGAACC 6446  
Qy 2401 TGGCGCTGCAAGCGGCAAGCAAGCGCTCTCTGACCAACCTTCCGACGAGCGGCA 2460  
Db 6447 TGGCGCTGCAAGCGGCAAGCAAGCGCTCTCTGACCAAGCGCTTCCGACGAGCGGCA 6506  
Qy 2461 GAGAGTCAAGCTTGAAGTGTATCTCTCAATGCCCCCTTGAAGGAGGAGCGCGGAGAC 2520  
Db 6507 GAGAGTCAAGCTTGAAGTGTATCTCTCAATGCCCCCTTGAAGGAGGAGCGCGGAGAT 6566

```

Qy 2521 CCCGATCTCAGCAGCGGGTCTTGTCTACCGGTAGTGAGAGCGCGGTGAGACGTCGTC 2580
Db 6567 CCCGATCTCAGCAGCGGGTCTTGTCTACCGGTAGTGAGAGCGCGGTGAGACGTCGTC 6626
Qy 2581 TGTGCTCGATGTCCTACATGATGAGAGCGCGCTCTGTATCAAGCCATCGCTCGAGAGAA 2640
Db 6627 TGTGCTCGATGTCCTACATGATGAGAGCGCGCTCTGTATCAAGCCATCGCTCGAGAGAA 6686
Qy 2641 AGCAGCTGCGCATCAACCGGCTTGAACGACCTCTTGTGTGTCACCAACATGTCCTAC 2700
Db 6687 ACCAGCTGCGCATCAACGACCTGAGACCTCTTGTGTGTCACCAACATGTCCTAT 6746
Qy 2701 GCTACACATCCCGCAGCGCAGACCGCAGAGAGGTCACTTTGACAGACTGCA 2760
Db 6747 GCTACACATCTCGCAGCGCAGACCGCTCGCAGAGAGGTCACTTTGACAGACTGCA 6806
Qy 2761 ATCTGAGACGATCACTACAGAGAGCTGCTCAGAGAGATGAGAGCGGCTCCACAGTT 2820
Db 6807 GTCCTGAGACGATCACTACCGAGAGCTGCTCAGAGAGATGAGAGCGGCTCCACAGTT 6866
Qy 2821 AAGGCTAAGCTTATCACTAGTAGAGAGAGCTGCAAGTCAAGCCCACTTTGGCCAA 2880
Db 6867 AAGGCTAAGCTTATCACTAGTAGAGAGAGCTGCAAGTCAAGCCCACTTTGGCCAA 6926
Qy 2881 TCTAATTTGGCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
Db 6927 TCTAATTTGGCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6986
Qy 2941 ATCCGCTCCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
Db 6987 ATCCGCTCCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7046
Qy 3001 ATGCGCAAAAATAGAGTTTCTGCTCCACCAAGAGAGAGAGAGAGAGAGAGAGAG 3060
Db 7047 ATGCGCAAAAATAGAGTTTCTGCTCCACCAAGAGAGAGAGAGAGAGAGAGAGAG 7106
Qy 3061 CTATGCTGTTCCAGAGCTTGGGGGTCCTGCTGTGAGAGAGAGAGAGAGAGAGAGAG 3120
Db 7107 CTATGCTGTTCCAGAGCTTGGGGGTCCTGCTGTGAGAGAGAGAGAGAGAGAGAGAG 7166
Qy 3121 GTCTCAGACCTCCCTCAGAGCTGTGATGAGCTCTCTGATCGAATTCAGTATCTCTG 3180
Db 7167 GTCTCAGACCTCCCTCAGAGCTGTGATGAGCTCTCTGATCGAATTCAGTATCTCTG 7226
Qy 3181 CAGCGGGTGAAGTCTCTGTGAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 3240
Db 7227 CAGCGGGTGAAGTCTCTGTGAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 7286
Qy 3241 TATGACACCGCTGTTTGAAGTCAACAGTCACTGAGATGACATCCGTGATGAGAGATCA 3300
Db 7287 TATGACACCGCTGTTTGAAGTCAACAGTCACTGAGATGACATCCGTGATGAGAGATCA 7346
Qy 3301 ATTATCAATGTTGTGAGCTTGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
Db 7347 ATCTACCAATGTTGTGAGCTTGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7406
Qy 3361 CGGCTTATATCGGGGCTCCCTGATCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
Db 7407 CGGCTTATATCGGGGCTCCCTGATCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 7466
Qy 3421 TGCAGGCGAGAGCGGCTGTGAGAGAGTGTGAGATGAGATGAGATGAGATGAGATGAG 3480
Db 7467 TGCAGGCGAGAGCGGCTGTGAGAGAGTGTGAGATGAGATGAGATGAGATGAGATGAG 7526
Qy 3481 GCGCTGAGAGCTGTGAGAGTGTGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 3540
Db 7527 GCGCTGAGAGCTGTGAGAGTGTGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 7586
Qy 3541 GGGCTTGTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Db 7587 GACCTTGTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7646
Qy 3601 TTCACGAGGCTATGACTAGTACTGTGCCCCCGGGGAGAGAGAGAGAGAGAGAGATAC 3660

```

```

Db 7647 TTCACGAGGCTATGACTAGTACTGTGCCCCCGGGGAGAGAGAGAGAGAGAGATAC 7706
Qy 3661 GACCTGAGTGTATATCATCATGCTCTCTCATATGTGTGCTGCGCAGCATGATCTGGC 3720
Db 7707 GACCTGAGTGTATATCATCATGCTCTCTCATATGTGTGCTGCGCAGCATGATCTGGC 7766
Qy 3721 AAAAGGTAATACTCAACCGGTGACCC 3749
Db 7767 AAAAGGTAATACTCAACCGGTGACCC 7795

RESULT 12
AAA98969
ID AAA98969 standard; DNA; 8649 BP.
XX
AC AAA98969;
XX
DT 08-FEB-2001 (first entry)
XX
DE Hepatitis C virus DNA fragment SEQ ID NO: 5.
XX
OS Hepatitis C virus.
XX
KW Cell culture; therapy; infection; vaccine; diagnosis; gene therapy; ds.
XX
PN DE19915178-A1.
XX
PD 05-OCT-2000.
XX
PF 03-APR-1999; 99DE-01015178.
XX
PR 03-APR-1999; 99DE-01015178.
XX
PA (UYMA-) UNIV MAINZ GUTENBERG JOHANNES.
XX
PI Bartschlagel R;
XX
DR WPI; 2000-629140/61.
XX
PT Cell culture system for hepatitis C virus, useful e.g. in screening for
PT therapeutic agents, comprises human hepatoma cells containing a viral RNA
PT construct that includes a selectable gene.
XX
PS Claim 9; Page 44-50; 58bp; German.
XX
CC This invention describes a novel Hepatitis C virus (HCV) cell culture
CC system comprising human hepatoma cells that contain an integrated HCV-RNA
CC construct (I). (I) contains the HCV-specific RNA segments 5'-NTR (non-
CC translated region), NS (non-structural)3, NS4A, NS4B, NS5A, NS5B and 3'-
CC NTR, and a selectable (marker) gene (II). The cell cultures, and/or (I),
CC are used to prepare, evaluate and/or test therapeutic and/or diagnostic
CC agents for HCV infections, and to prepare vaccines against HCV infection
CC (particularly preparation of attenuated HCV). The can also be used for
CC preparation of a liver-specific delivery system for gene therapy, and to
CC identify cells permissive for HCV replication. Virus RNA replicates
CC autonomously and with high efficiency in this cellular system, so that
CC variations in replication rates can be measured (for screening antiviral
CC agents) quantitatively or qualitatively, using standard laboratory
CC equipment. Efficient replication of HCV RNA is only achieved when the
CC specified RNA segments are present and when the transfected cells are
CC maintained under permanent selection pressure
XX
SQ Sequence 8649 BP; 1773 A; 2563 C; 2437 G; 1876 T; 0 U; 0 Other;

Query Match 87.8%; Score 3291.4; DB 3; Length 8649;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

Qy 1 TGGAGAGGCGTCTTACAGGCTCAACCCAGTGTATGCCACTTCTGTCCAAACAAG 60
Db 4057 TGGAGAGGCGTCTTACAGGCTCAACCCAGTGTATGCCACTTCTGTGTCCAGACTAAG 4116

```

QY 61 CAGGAGAGAGAACTTCCCTTACTGTGTGGCTACAGGCTACTGTGTGCGCTAGGACC 120  
DB 4117 CAGGAGAGAGAACTTCCCTTACTGTGTAGCATACAGGCTACGGTGTGCGCAGAGGCT 4176  
QY 121 CAGGCCCACTCTCATCATGTGGATCAAAATGTGAAGTGTCTCATACGGCTTAAAGCTTACT 180  
DB 4177 CAGGCTCCACTTCATCTGTGGAGCAAAATGTGAAGTGTCTCATACGGCTTAAAGCTTACT 4236  
QY 181 CTGGGGGGGCAACACCTTGTGTATPAGGCTGGGAGCGGCTCCAAAACGAGCTCACCTTC 240  
DB 4237 CTGGAGGGGCAAGCGCCCTGTGTATPAGGCTGGGAGCGGCTCCAAAACGAGCTTACTTACC 4296  
QY 241 ACACACCCCATTAACAAATTCATCATGTGCATGTGATGACGCGACCTGAGAGTCTGTACG 300  
DB 4297 ACACACCCCATTAACAAATTCATCATGTGCATGTGATGACGCGCTGACCTGGAGGTCGTACG 4356  
QY 301 AGCAGCTGGGTGTGTGGGGGGGCTCTTGGAGCTGTGGCTGTGCTGTGCTTGTGACACAC 360  
DB 4357 AGCAGCTGGGTGTGTGGGGGGGCTCTTGGAGCTGTGGCTGTGCTGTGCTTGTGACACAC 4416  
QY 361 GGGAGGTGTGTATGTGGGTAGATCATGTGTCCGGGGGGCGGCTATTGTTCGCGAC 420  
DB 4417 GGGAGGTGTGTATGTGGGTAGATCATGTGTCCGAAAAGCGGCTCATTTCCCGAC 4476  
QY 421 AGGGAAGTCTCTACAGAGAGTTCAGATGAGATGAAAGTGCAGCTGTGCACTTCCCTTAC 480  
DB 4477 AGGGAAGTCTCTTACGGGGAGTTGATGAGATGAAAGTGGGCTGTACACTTCCCTTAC 4536  
QY 481 ATGAGAGAGGAATGCACTGTCCGAGCACTTCAAGCAAAAGCGCTCGGTTGTGTGACG 540  
DB 4537 ATGGAACAGGGAATGCACTGTCCGCAACAAATTCAAAACAGAGCAATCGGTTGTGTGCA 4596  
QY 541 ACAGCAACCAAGAGAGGGAGGGGCTCCGCTGGGTGGAGTCCAGGTGGAGCGGCTT 600  
DB 4597 ACAGCAACCAAGAGAGGGAGGGGCTCCGCTGGGTGGAGTCCAGGTGGAGCGGCTT 4656  
QY 601 GAGACCTTCTGGGCGAAACCATGTGGAACTTATCATGCGGGATACAGTATAGACAGC 660  
DB 4657 GAGACCTTCTGGGCGAAACCATGTGGAACTTATCATGCGGGATACAGTATAGACAGC 4716  
QY 661 TTGTTCATCTGTCTGGGAAATCCCGGATGTGATCACTGATGGCTTACAGCTCTGTCT 720  
DB 4717 TTGTTCATCTGTCTGGGAAATCCCGGATGTGATCACTGATGGCTTACAGCTCTGTCT 4776  
QY 721 ACTAGCCGCTCAACCAATGATCCCTGTGTTAACTCCGGGGGGATGGGTAGCC 780  
DB 4777 ACTAGCCGCTCAACCAATGATCCCTGTGTTAACTCCGGGGGGATGGGTAGCC 4836  
QY 781 GCCCACTCGCTCCCGGAGTGTGCTTCACTTTCGTAGGCGCGGCAATTGTCTGTGCG 840  
DB 4837 GCCCACTCGCTCCCGGAGTGTGCTTTCGTAGGCGCGGCAATTGTCTGTGCG 4896  
QY 841 GCTGTGTGAGCATAGGCTTGGGAGGTGTGTGATCTTGGCGGGCTATGAGCA 900  
DB 4897 GCTGTGTGAGCATAGGCTTGGGAGGTGTGTGATCTTGGCGGGCTATGAGCA 4956  
QY 901 GAGGTGGCAGGCGGCTGTGGCTTAAAGTCAATGAGGAGGAGAAATGCCCTCACCGAG 960  
DB 4957 GAGGTGGCAGGCGGCTGTGGCTTAAAGTCAATGAGGAGGAGAAATGCCCTCACCGAG 5016  
QY 961 GACCTGTAACTTACTCTCCGTGCATCTCTCTCTGTGTGTCCTGTGTCTGTGGAGTCTG 1020  
DB 5017 GACCTGTAACTTACTCTCCGTGCATCTCTCTCTGTGTGTCCTGTGTCTGTGGAGTCTG 5076  
QY 1021 TGGCAGCGATACTGTGCGTGTGCACTGTGGTTCAGAGGAGGAGGCTGTGCACTGTGATGAC 1080  
DB 5077 TGGCAGCGATACTGTGCGTGTGCACTGTGGTTCAGAGGAGGAGGCTGTGCACTGTGATGAC 5136  
QY 1081 CGGCTGATAGCTTGTGCGCTGTGGGGGTAACTGTTTCCCGCAGCATATGTGTGCGAG 1140  
DB 5137 CGGCTGATAGCTTGTGCGCTGTGGGGGTAACTGTTTCCCGCAGCATATGTGTGCGAG 5196  
QY 1141 AGGAGCGCGAGACGTGTCACTCAGATCTCTCCGACCTTACTATCACCAACTGTGTG 1200

DB 5197 AGCGACGCTGACGACAGTGTCACTCAGATCCCTCTAAGTCTTACATCACTACGCTGCTG 5256  
QY 1201 AAGAGGCTCCACAGTGAATTAAGAGGACGTGTCCACGCGCTGTCCGGCTGTGGCTTA 1260  
DB 5257 AAGAGGCTCCACAGTGAATTAAGAGGACGTGTCCACGCGCTGTCCGGCTGTGGCTTA 5316  
QY 1261 AAGGATGTTTGGAGCTGATATGACAGTGTGTGTGACTTCAAGACCTGTGCTCAGTTC 1320  
DB 5317 AAGATGTTTGGAGTGTGATATGACAGTGTGTGACTTCAAGACCTGTGCTCAGTTC 5376  
QY 1321 AAGCTCTGTGCGGATTAACGGGAGTCCCTTTTCTCATGTCCAACTGTGGTGAAGAGG 1380  
DB 5377 AAGCTCTGTGCGGATTAACGGGAGTCCCTTTTCTCATGTCCAACTGTGGTGAAGAGG 5436  
QY 1381 GTGTGGGGGGAGAGGACATGATGAGACCACTGTCTATGTGGAGACAGATCACCGGA 1440  
DB 5437 GTGTGGGGGGAGAGGACATGATGAGACCACTGTCTATGTGGAGACAGATCACCGGA 5496  
QY 1441 CATGTCAAAAACGTTTCCATGAGATCTGTGGGCTTAAGACCTGTATGATATGTGCAAT 1500  
DB 5497 CATGTCAAAAACGTTTCCATGAGATCTGTGGGCTTAAGACCTGTATGATATGTGCAAT 5556  
QY 1501 GGAACATTCCTCATACGATATACACAGGCGCTGTGACGCGCTTCCCGACGCCAATC 1560  
DB 5557 GGAACATTCCTCATACGATATACACAGGCGCGCTGTGACGCGCTTCCCGACGCCAATC 5616  
QY 1561 TATTCAGGGGCTGTGTGGGGGTGTGGGTGTGAGAGTATGAGAGGTTACGGGGGTGGG 1620  
DB 5617 TATTCAGGGGCTGTGTGGGGGTGTGGGTGTGAGAGTATGAGAGGTTACGGGGGTGGG 5676  
QY 1621 GATTTTCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCGCGGTCCAGGTTTCA 1680  
DB 5677 GATTTTCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCGCGGTTCAGGTTTCCG 5736  
QY 1681 GCCCGGAAATTTTTCACAGAAATGATGGGGGTGCGGCTGTGACAGTATCCCTCGGCTGT 1740  
DB 5737 GCCCGGAAATTTTTCACAGAAATGATGGGGGTGCGGCTGTGACAGTATCCCTCGGCTGT 5796  
QY 1741 AAAACCTCTCTACGGGAGAGGTTCATTCAGAGTGTGGGCTTCAACCAATACCTGGTTGG 1800  
DB 5797 AAAACCTCTCTACGGGAGAGGTTCATTCAGAGTGTGGGCTTCAACCAATACCTGGTTGG 5856  
QY 1801 TGCAGGCTCCATGCGAGCGCCGACCGAGTGTAGACGTGTCACTTCACTGTCTACCGAC 1860  
DB 5857 TGCAGGCTCCATGCGAGCGCCGACCGAGTGTAGACGTGTCTCATGTCTACCGAC 5916  
QY 1861 CCTTCCCATATACAGCAGAGACGGTTAAGCGAGGCTGTGCGAGGGGTCTCCCTCTTC 1920  
DB 5917 CCTTCCCATATACAGCAGAGACGGTTAAGCGAGGCTGTGCGAGGGGTCTCCCTCTTC 5976  
QY 1921 TTGGCAGGCTTTCAGCTAGCCAGTGTGTGGGCTTCTCTGAGGCGCATTAATTTAC 1980  
DB 5977 TTGGCAGGCTTTCAGCTAGCCAGTGTGTGGGCTTCTCTGAGGCGCATTAATTTAC 6036  
QY 1981 CAAATGATCTTCCAGACGCTGACCTCATGAGGCGCACTCTGTGTGGCGGATGAGATG 2040  
DB 6037 CAGTATGATCTTCCCGAGCGCTGACCTCATGAGGCGCACTCTGTGTGGCGGAGATG 6096  
QY 2041 GGGGGGACATTAACCGGCTGTGATGAGAAACAAAGTATGTAATCTGTGACTTTTGAC 2100  
DB 6097 GGGGGGACATTAACCGGCTGTGATGAGAAACAAAGTATGTAATCTGTGACTTTTGAC 6156  
QY 2101 CGGCTCCGAGCGGAGAGATGAGCGGGGAAAGTCCGTCGGGCGGAGATTCGCGGAAA 2160  
DB 6157 CGGCTCCGAGCGGAGAGATGAGCGGGGAAAGTCCGTCGGGCGGAGATTCGCGGAAA 6216  
QY 2161 TCCAGAAATTTCCACACAGATGCGCGCATGGGCAAGCGCGGATTAACACCTTCCGCTG 2220  
DB 6217 TCCAGAAATTTCCCTGAGCGATGCGCATATGGGCAAGCGCGGATTAACACCTTCCACTG 6276  
QY 2221 CTGGAATCTTGAAGGCGCGGACATAGTCTCTTCACTGTGTATCATGGGTGCCACTGTGCA 2280

Db 6277 TTAGAGTCTGGAAGGACCGGACCTAGTCCCTCCAGTGTAGACGGGTGTCATTCGCC 6336  
 Qy 2281 CCGACTAAGACCCCTCCTATACACACTCCACGGAGGAGAGGACAGTTGTTGACAGAA 2340  
 Db 6337 CCGCAGAGGCCCCCTCGATACACCTCCAGAGGAGAGGAGCGTTGTCCTGTCCGAA 6396  
 Qy 2341 TCCACCGTGTCTTCTGCTCGGAGAGCTTGCCACAAAGGCTTTCGGTAGCTCCGAAACG 2400  
 Db 6397 TCTACCGTGTCTTCTGCTCGGAGAGCTTGCCACAAAGGCTTTCGGTAGCTCCGAAACG 6456  
 Qy 2401 TCGCCGCTGACAGCGGACCGGACCGGCTCTGACCAACCTTCGACGACGCGGAC 2460  
 Db 6457 TCGCCGCTGACAGCGGACCGGACCGGCTCTGACCAACCTTCGACGACGCGGAC 6516  
 Qy 2461 GAGAGTCTGACGCTGATGCTATTCCTCCATGCCCCCTTACGAGGAGGAGCGGAGGAC 2520  
 Db 6517 GCGAGATCCGACCTTGAAGTGTACTCTCCATGCCCCCTTACGAGGAGGAGCGGAGGAC 6576  
 Qy 2521 CCGGATCTGACGACCGGCTTGTGCTACCGTAGTAGAGGAGCGGCTGAGACGTCGTC 2580  
 Db 6577 CCGGATCTGACGACCGGCTTGTGCTACCGTAGTAGAGGAGCTGAGACGTCGTC 6636  
 Qy 2581 TGTGCTGATGCTCTACACATGAGACAGGCGCTGTATCAAGCATGCGCTGCGAGAA 2640  
 Db 6637 TGTGCTGATGCTCTACACATGAGACAGGCGCTGTATCAAGCATGCGCTGCGAGAA 6696  
 Qy 2641 AGCAAGTCCCATCAACGCGTTGAGCACTCTTGTGCTGCTACCAACATGCTCTAC 2700  
 Db 6697 ACCCAAGTCCCATCAACGCGTTGAGCACTCTTGTGCTGCTACCAACATGCTCTAC 6756  
 Qy 2701 GCTACACATCCCGACGCGCAACGCGGACGAGAGGCTCACTTTGACAGATGCA 2760  
 Db 6757 GCTACACATCCCGACGCGCAACGCGGACGAGAGGCTCACTTTGACAGATGCA 6816  
 Qy 2761 ATCTTGAGAGTCACTACAGAGCGTGTCAAGAGATGAGAGGCGTCCACAGTT 2820  
 Db 6817 GTCTTGAGAGTCACTACAGAGCGTGTCAAGAGATGAGAGGCGTCCACAGTT 6876  
 Qy 2821 AAGGCTAAGCTTCTATCAGTAGAGAGGCTGCAAGTCAAGGCTCCCACTTTGGGCCAA 2880  
 Db 6877 AAGGCTAAGCTTCTATCAGTAGAGAGGCTGCAAGTCAAGGCTCCCACTTTGGGCCAA 6936  
 Qy 2881 TCTAATTTGGCTATGGGCAAGAGCGTCCGAGACCTTACGACAGGCTTTAACAC 2940  
 Db 6937 TCTAATTTGGCTATGGGCAAGAGCGTCCGAGACCTTACGACAGGCTTTAACAC 6996  
 Qy 2941 ATCCGCTCGGTGGAGGACTTGTGAGAGACCTGAACCAATTGACACCAACATC 3000  
 Db 6997 ATCCGCTCGGTGGAGGACTTGTGAGAGACCTGAACCAATTGACACCAACATC 7056  
 Qy 3001 ATGGCAAAAATGAGGTTTTTCTGCTCCACAGAGAGAGGAGCGGCAAGCCAGCTCCG 3060  
 Db 7057 ATGGCAAAAATGAGGTTTTTCTGCTCCACAGAGAGAGGAGCGGCAAGCCAGCTCCG 7116  
 Qy 3061 CTTATGCTGTTCCAGACTTGGGGGTCGCTGTGAGAGAAATGGCCCTTATGACGAG 3120  
 Db 7117 CTTATGCTGTTCCAGACTTGGGGGTCGCTGTGAGAGAAATGGCCCTTATGACGAG 7176  
 Qy 3121 GTCTCCACCTCCCTCAGGCTGATGAGCTCCTGTAAGATTCAGATTCTCCTGGA 3180  
 Db 7177 GTCTCCACCTCCCTCAGGCTGATGAGCTCCTGTAAGATTCAGATTCTCCTGGA 7236  
 Qy 3181 CAGCGGCTGAGTCTGCTGTAAGAGCTGGAATCAAAAGAGCCCTTATGGCTTTGCA 3240  
 Db 7237 CAGCGGCTGAGTCTGCTGTAAGAGCTGGAATCAAAAGAGCCCTTATGGCTTTGCA 7296  
 Qy 3241 TATGACACCGGCTGTTTGAAGTCAAGTCACTGAGATGACATCCGTGTAAGAGGATCA 3300  
 Db 7297 TATGACACCGGCTGTTTGAAGTCAAGTCACTGAGATGACATCCGTGTAAGAGGATCA 7356  
 Qy 3301 ATTATCAATGTTGACTTGGCCCCGAGAGCCAGACAGGCTTAAGTCTGCTACAGAG 3360  
 Db 7357 ATTATCAATGTTGACTTGGCCCCGAGAGCCAGACAGGCTTAAGTCTGCTACAGAG 7416

Qy 3361 CGGCTTATATCGGGGGTCCCTGACTAATTAAGAGGAGAGAACTGGGCTATCGCCGG 3420  
 Db 7417 CGGCTTATATCGGGGGTCCCTGACTAATTAAGAGGAGAGAACTGGGCTATCGCCGG 7476  
 Qy 3421 TGCAGGAGGAGGCTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3480  
 Db 7477 TGCAGGAGGAGGCTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7536  
 Qy 3481 GCTCTGACGCTGTGTGAGCTGCAAAAGTTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 3540  
 Db 7537 GCTCTGACGCTGTGTGAGCTGCAAAAGTTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 7596  
 Qy 3541 GCGCTGTGCTATCTGTGAGAGCGGAGGAGACCGAGAGAGAGCGGCGGAGCTTACGAGTC 3600  
 Db 7597 GACCTGTGCTATCTGTGAGAGCGGAGGAGACCGAGAGAGAGCGGCGGAGCTTACGAGTC 7656  
 Qy 3601 TTCACGAGGCTATGACTAGTATCTGCCCCCGGAGGAGCGGCGGAGCTTACGAGTC 3660  
 Db 7657 TTCACGAGGCTATGACTAGTATCTGCCCCCGGAGGAGCGGCGGAGCTTACGAGTC 7716  
 Qy 3661 GACCTGAGTGAATACATCATGCTCTCTCAATGTGTGCTGCGGACGATGATCTGGC 3720  
 Db 7717 GACCTGAGTGAATACATCATGCTCTCTCAATGTGTGCTGCGGACGATGATCTGGC 7776  
 Qy 3721 AAAAGGTAATACCTCAACCGGTGACC 3749  
 Db 7777 AAAAGGTAATACCTCAACCGGTGACC 7805

RESULT 13  
 ABK91424  
 ID ABK91424 standard; DNA; 9605 BP.  
 AC ABK91424;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE Hepatitis C virus Con 1 isolate DNA mutant 1.  
 XX  
 XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
 OS Hepatitis C virus.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 342..9374  
 FT /tag= a  
 FT /product= "HCV polyprotein"  
 FT /note= "The polyprotein consists of the Core, E1, E2, P7,  
 FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
 FT mutation replace(3625,G)  
 FT /tag= b  
 PN MO2002S9321-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 16-JAN-2002; 2002MO-BP000526.  
 XX  
 PR 23-JAN-2001; 2001US-0263479P.  
 XX  
 PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.  
 XX  
 PI De Francesco R, Migliaccio G, Paonessa G;  
 DR WPI; 2002-599793/64.  
 XX  
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and

PT expression.

XX Claim 9; Page: 69pp; English.

XX The invention relates to nucleic acid molecules comprising altered HCV  
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
CC internal ribosome entry site (IRES) region coding for one or more NS3,  
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
CC are detailed in the specification. Also included are (1) an expression  
CC vector comprising a nucleotide sequence coding for the altered nucleic  
CC acid, which is transcriptionally coupled to an exogenous promoter; (2) a  
CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC proteins, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is an HCV replicon. Con 1  
CC mutant of the invention. Note: The present sequence is not shown in the  
CC specification but was created by the indexer using the HCV sequence  
CC appearing as ABK91411 and the information in Claim 9

XX Sequence 9605 BP; 1910 A; 2884 C; 2732 G; 2079 T; 0 U; 0 Other;

Query Match 87.8%; Score 3291.4; DB 6; Length 9605;

Best Local Similarity 92.4%; Pred. No. 0;

Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 1 TGGGAGGGGCGTCTTCAAGGCGCTCACCCAGTGGAGTGGCCACTTCCGTCCCAAGAAAG 60  
DB 5013 TGGGAGGCGTCTTCAAGGCGCTCACCCAGTGGAGTGGCCACTTCCGTCCCAAGAAAG 5072  
QY 61 CAGGACGAGAGACAATCTCCCTACCTGTGCGGTACAGGCTACTGTGTGCGGTAGAGCC 120  
DB 5073 CAGGACGAGAGACAATCTCCCTACCTGTGCGGTACAGGCTACTGTGTGCGGTAGAGCC 5132  
QY 121 CAGGCGCCACCTCCATCATATGGGATCAAAATGTGGAAGTGTCTCATACGGCTAAAGCTACT 180  
DB 5133 CAGGCTTCACCTCCATCATATGGGATCAAAATGTGGAAGTGTCTCATACGGCTAAAGCTACT 5192  
QY 181 CTGCGGGGGGCAACCCCTGTGTATAGGCTGGGAGCGCTCCAAACGAGGTACCCCTC 240  
DB 5193 CTGCGGGGGGCAACCCCTGTGTATAGGCTGGGAGCGCTCCAAACGAGGTACCCCTC 5252  
QY 241 ACAACACCCCATTAACAATTCATCATATGCGATGATGACGACCTGAGAGTGTACAG 300  
DB 5253 ACAACACCCCATTAACAATTCATCATATGCGATGATGACGACCTGAGAGTGTACAG 5312  
QY 301 AGCACTGGGTGTGTGTGGCGGGGCTCTTGCAAGCTCTGGCTGTGATTGCTTGACACA 360  
DB 5313 AGCACTGGGTGTGTGTGGCGGGGCTCTTGCAAGCTCTGGCTGTGATTGCTTGACACA 5372  
QY 361 GGCAGCGTGTGCTATGTGGGTAGAGATCATCTGTGCGGGGCGCGGTATTTGTTCCGAC 420  
DB 5373 GGCAGCGTGTGCTATGTGGGTAGAGATCATCTGTGCGGGGCGCGGTATTTGTTCCGAC 5432  
QY 421 AGGGAAGTCTCTTACAGAGATTGATGATGAGAAAGTGCAGTCTCCCTTAC 480  
DB 5433 AGGGAAGTCTCTTACAGAGATTGATGATGAGAAAGTGCAGTCTCCCTTAC 5492  
QY 481 ATGAGAGAGGAATGAGCTTCGCCGAGCAGTTCAGCAAGCAAAAGCGCTCGGCTGTGACG 540  
DB 5493 ATGAGAGAGGAATGAGCTTCGCCGAGCAGTTCAGCAAGCAAAAGCGCTCGGCTGTGACG 5552  
QY 541 ACAGCCCAACGAAGGAGGAGCGCTGTCTCCGTGTGAGATGCAAGTGGCGAGCCCTT 600  
DB 5553 ACAGCCCAACGAAGGAGGAGCGCTGTCTCCGTGTGAGATGCAAGTGGCGAGCCCTT 5612

QY 601 GAGACCTTCTGGGCGAAGACATGTGGAACTTCATCAGCGGAGTACAGTACTTACAGGC 660  
DB 5613 GAGACCTTCTGGGCGAAGACATGTGGAACTTCATCAGCGGAGTACAGTACTTACAGGC 5672  
QY 661 TTGTCCACTCTGCGTGGGAAATCCCGGATGTGATCATGATGCGCTTCCACACCTCTGTC 720  
DB 5673 TTGTCCACTCTGCGTGGGAAATCCCGGATGTGATCATGATGCGCTTCCACACCTCTTATC 5732  
QY 721 ACTAGCCGCTCACCACCAACCAATCTACCCCTGCTTAAATCATCTGCGGGGAGTGGTAGCC 780  
DB 5733 ACTAGCCGCTCACCACCAACCAATCTACCCCTGCTTAAATCATCTGCGGGGAGTGGTAGCC 5792  
QY 781 GCCCACTGCTCCCCAGTCTGCTTCAAGCTTTGTAAGCGCCGCACTTCTGTGTCG 840  
DB 5793 GCCCACTGCTCCCCAGTCTGCTTCAAGCTTTGTAAGCGCCGCACTTCTGTGTCG 5852  
QY 841 GCTTTTGGAGATAGGCGCTTGGGAAAGGTCTTGTGAAATCTTGGCGGGGCTATGGAGA 900  
DB 5853 GCTTTTGGAGATAGGCGCTTGGGAAAGGTCTTGTGAAATCTTGGCGGGGCTATGGAGA 5912  
QY 901 GGAGTGGCAGGCGCGCTCGTGGCTTTTAAAGTCAATGAGCGGCAAAATGCTTCCACCGAG 960  
DB 5913 GGAGTGGCAGGCGCGCTCGTGGCTTTTAAAGTCAATGAGCGGCAAAATGCTTCCACCGAG 5972  
QY 961 GACCTGTTAACTTACTCTCCCTGCATCTCTCTGTCGCTGTGCTGTGCGGCTGTG 1020  
DB 5973 GACCTGTTAACTTACTCTCCCTGCATCTCTCTGTCGCTGTGCTGTGCGGCTGTG 6032  
QY 1021 TGGCAGCGCATCTGCTGTGGGACATGCGGTCCAGGGGAGGGGCTGTGAGATGATGAA 1080  
DB 6033 TGGCAGCGCATCTGCTGTGGGACATGCGGTCCAGGGGAGGGGCTGTGAGATGATGAA 6092  
QY 1081 CGGCTGATAGCGTTCGCTCGCGGGGTAACATGTTTCCCGACGACATATGCGCAGAG 1140  
DB 6093 CGGCTGATAGCGTTCGCTCGCGGGGTAACATGTTTCCCGACGACATATGCGCTAG 6152  
QY 1141 AGCAGCGCGCAGCAGTGTCACTGATCTCTCTCCGACTTATCAACCACTGTTG 1200  
DB 6153 AGCAGCGCGCAGCAGTGTCACTGATCTCTCTCCGACTTATCAACCACTGAGCTGTG 6212  
QY 1201 AAGAGCTCCACCACTGATGATTAACAGAGACATGCTCCAGCCCTGCTCGGGCTGTG 1260  
DB 6213 AAGAGCTCCACCACTGATGATTAACAGAGACATGCTCCAGCCCTGCTCGGGCTGTG 6272  
QY 1261 AAGGATGTTTGGGATCGATATGACAGATTTTGGCTGATTAAGACCTGCTCCAGTCC 1320  
DB 6273 AAGGATGTTTGGGATCGATATGACAGATTTTGGCTGATTAAGACCTGCTCCAGTCC 6332  
QY 6333 AAGCTCTGCGCGATTTACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTA 1380  
DB 6333 AAGCTCTGCGCGATTTACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTA 6392  
QY 1381 GTCTGGCGGGGAGACGGGATCATGACAGACATCTGTCTATGTGAGCAAGATCCGGA 1440  
DB 6393 GTCTGGCGGGGAGACGGGATCATGACAGACATCTGTCTATGTGAGCAAGATCCGGA 6452  
QY 1441 CATGTCAAAAACGTTCCATGAGGATCGTTGGGCTTAAGACCTGTATGATTAAGTGGCAT 1500  
DB 6453 CATGTCAAAAACGTTCCATGAGGATCGTTGGGCTTAAGACCTGTATGATTAAGTGGCAT 6512  
QY 1501 GGAACATTTCCCATTAACCGGTAACCAAGGAGCCCTGTGACGCGCTCCAGCGCCAAC 1560  
DB 6513 GGAACATTTCCCATTAACCGGTAACCAAGGAGCCCTGTGACGCGCTCCAGCGCCAAC 6572  
QY 1561 TATTTCAGGGCGCTGTGGGGGTGCTGCTGAGAGATACGTGAGGTTACGGGGGTGGG 1620  
DB 6573 TATTTCAGGGCGCTGTGGGGGTGCTGCTGAGAGATACGTGAGGTTACGGGGGTGGG 6632  
QY 1621 GATTTCACATAGTGAAGGATGACCTGACCAAGTAAATTTGCGCGCAGGTTTCA 1680  
DB 6633 GATTTCACATAGTGAAGGATGACCTGACCAAGTAAATTTGCGCGCAGGTTTCCG 6692  
QY 1681 GCCCCGAATTTCTTCAAGAAATGATGAGGAGTGCAGGCTGACAGGTAAGCTCCGCGTGC 1740

Db 6693 GCGCCGGAATTCCTTCAAGAGGATGAGGAGTGGGCTGGGTGACAGATACGCTCCAGCGTGC 6752  
Qy 1741 AAAACCTCTCTACGGGAGAGGATCAATTCAGTGGGCTCAACCAATTCCTGTTGGG 1800  
Db 6753 AAAACCTCTCTACGGGAGAGGATCAATTCCTGTTGGGCTCAATTCCTGTTGGG 6812  
Qy 1801 TCGAGACTCCATGACGAGCCGGAACGGATGTAGAGTGTCACTTCCATGCTCACGAC 1860  
Db 6813 TCACAGCTCCATGACGAGCCGGAACGGATGTAGAGTGTCACTTCCATGCTCACGAC 6872  
Qy 1861 CCCTCCCATCAACAGCAAGACGAGTAAAGCGAGGCTGGCCAGGAGGCTTCCCTCC 1920  
Db 6873 CCTCCCATTAACGAGGAGAGCGGCTAAGCTAGGCTGGCCAGGAGATCTCCCTCC 6932  
Qy 1921 TTGGCAGCTTTCACTAGTCCAGTGTCTGCGCTTCTCCAGAGGCTCAATCATTTAC 1980  
Db 6933 TTGGCAGCTTTCACTAGTCCAGTGTCTGCGCTTCTCCAGAGGCTCAATCATTTAC 6992  
Qy 1981 CAAATGACTTCCAGAGCTGACCTCATGAGGACCACTCTGTGGCGGATGAGATG 2040  
Db 6993 CATTGATCTCCCGGAGCGCTGACCTCATGAGGACCACTCTGTGGCGGAGAGATG 7052  
Qy 2041 GCGGAGGACATTAACCGCGTGAAGTCAAGAACAGATGTAATCTTGAATCTTTCAG 2100  
Db 7053 GCGGAGGACATTAACCGCGTGAAGTCAAGAACAGATGTAATCTTGAATCTTTCAG 7112  
Qy 2101 CCGCTCCGAGCGGAGAGATGAGCGGAGATCCGCTCCCGGAGATCTTCCGAGAA 2160  
Db 7113 CCGCTCCGAGCGGAGAGATGAGCGGAGATCCGCTCCCGGAGATCTTCCGAGAA 7172  
Qy 2161 TCCAGAAATTCACACAGAGATGCGGAGTGGGACGCGCGGATTAACAACCTTCGCTG 2220  
Db 7173 TCCAGAAATTCACACAGAGATGCGGAGTGGGACGCGCGGATTAACAACCTTCGCTG 7232  
Qy 2221 CTGAGATCTTGAAGGCGCGGACCTACGTCCTCCATGATGATGATGATGATGATGATG 2280  
Db 7233 TTAGAGTCTTGAAGGCGCGGACCTACGTCCTCCATGATGATGATGATGATGATG 7292  
Qy 2281 CCTACTTAAGACCTCTTATACCACTCCAGAGAGAGAGAGAGATGTTGTTGACAGAA 2340  
Db 7293 CCTGACAAAGGCGCGCTCCGATACCACTCCAGAGAGAGAGAGATGTTGTTGACAGAA 7352  
Qy 2341 TCCACGATGCTTCTGCGCTGCGGAGCTTGGCAAAAGCTTGGATGCTCCGAAACG 2400  
Db 7353 TCTACGATGCTTCTGCGCTTGGGAGCTGCGCAAAAGCTTGGATGCTCCGAAATG 7412  
Qy 2401 TCGGCGTGAAGCGGACGCGCAACGCGCTCTGACCAACCTTCCGACGAGCGGAG 2460  
Db 7413 TCGGCGTGAAGCGGACGCGCAACGCGCTCTGACCAACCTTCCGACGAGCGGAG 7472  
Qy 2461 GCAAGATCTGAGCGTGAAGTGTCTCTCCATGCCCCCTTGAAGGAGGAGCGGAGGAG 2520  
Db 7473 GCGGAGTCCGAGCGTGAAGTGTCTCTCCATGCCCCCTTGAAGGAGGAGCGGAGGAG 7532  
Qy 2521 CCGGATCTAGAGAGCGGCTCTGCTTACCGTGAAGAGAGCGCGTGAAGAGCGTCTG 2580  
Db 7533 CCGGATCTAGAGAGCGGCTCTGCTTACCGTGAAGAGAGCGCGTGAAGAGCGTCTG 7592  
Qy 2581 TCGCTGATGATCTTCAACATGAGAGAGCGCTCTGATCAGCGCATGCGTGGAGAGAA 2640  
Db 7593 TCGCTGATGATCTTCAACATGAGAGAGCGCGCTCTGATCAGCGCATGCGTGGAGAGAA 7652  
Qy 2641 AGCAAGCTGCCATCAACGCGTTGAGCAACTCTTCTGCGTCAACAACATGCTTAC 2700  
Db 7653 ACCAAGCTGCCATCAACGCGTTGAGCAACTCTTCTGCGTCAACAACATGCTTAC 7712  
Qy 2701 GGTACCAATCCCGAGCGGAGCGGACGCGGAGAGAGAGTCACTTGAAGAGTCAAG 2760  
Db 7713 GGTACCAATCCCGAGCGGAGCGGACGCGGAGAGAGAGTCACTTGAAGAGTCAAG 2772  
Qy 2761 ATCTGAGATCACTACGAGAGCGTGTCAAGAGATGAGGAGGAGCGTTCACAGTT 2820

Db 7773 GTCTTGAAGACCACTACCGGAGCGTGTCAAGAGATGAGGAGCGTTCACAGTT 7832  
Qy 2821 AAGGCTAAGCTTCTATCAGTGAAGAGCGTGAAGCGGCGGACCTTTCAGGAGAA 2880  
Db 7833 AAGGCTAAGCTTCTATCAGTGAAGAGCGTGAAGCGGCGGACCTTTCAGGAGAA 7892  
Qy 2881 TCTAATTTGGCTATGAGGAGAGAGAGCGTCCGGAACCTTATCCAGAGGAGCAATTAAC 2940  
Db 7893 TCTAATTTGGCTATGAGGAGAGAGAGCGTCCGGAACCTTATCCAGAGGAGCAATTAAC 7952  
Qy 2941 ATCCGCTCCGTGAGAGAGATTTGTTGAAGACCTGAACCAATTTGACACCAATC 3000  
Db 7953 ATCCGCTCCGTGAGAGAGATTTGTTGAAGACCTGAACCAATTTGACACCAATC 8012  
Qy 3001 ATGGCAAAATGAGGTTTTCGCGTCAACAGAGAGAGGAGCGGAGGAGCGTCCG 3060  
Db 8013 ATGGCAAAATGAGGTTTTCGCGTCAACAGAGAGAGGAGCGGAGGAGCGTCCG 8072  
Qy 3061 CTATGATGTTCCAGACTTGGGAGGCTCGTGTGTGAGAGAAATGAGGCTTATGACGTG 3120  
Db 8073 CTATGATGTTCCAGACTTGGGAGGCTCGTGTGTGAGAGAAATGAGGCTTATGACGTG 8132  
Qy 3121 GTCTCAACCTTCTCAGGCTGTGATGAGCTCTCTGATCGATTCAGATTTCTCTGGA 3180  
Db 8133 GTCTCAACCTTCTCAGGCTGTGATGAGCTCTCTGATCGATTCAGATTTCTCTGGA 8192  
Qy 3181 CAGCGGATGAGTTCCTGATGAGAGCGCTGGAATCAAGAGAGCGGCTTATGAG 3240  
Db 8193 CAGCGGATGAGTTCCTGATGAGAGCGCTGGAATCAAGAGAGCGGCTTATGAG 8252  
Qy 3241 TATGACACCGGCTGTTTACTCAACAGTCACTGAGATGATCCGTTAGAGAGTCA 3300  
Db 8253 TATGACACCGGCTGTTTACTCAACAGTCACTGAGATGATCCGTTAGAGAGTCA 8312  
Qy 3301 ATTTATCAATGTTGATGAGTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3360  
Db 8313 ATTTATCAATGTTGATGAGTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8372  
Qy 3361 CCGCTTATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3420  
Db 8373 CCGCTTATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8432  
Qy 3421 TCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3480  
Db 8433 TCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8492  
Qy 3481 GCGCTGAGCGCTGTGAGCTGCAAGAGTCCAGAGCTGCAAGAGTGTGTGCGGAGAG 3540  
Db 8493 GCGCTGAGCGCGCTGTGAGCTGCAAGAGTCCAGAGCTGCAAGAGTGTGTGCGGAGAG 8552  
Qy 3541 GCGCTGTGTGTATCTGTGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3600  
Db 8553 GACCTTGTGTGTATCTGTGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8612  
Qy 3601 TTCAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3660  
Db 8613 TTCAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8672  
Qy 3661 GACCTGAGTGTATTAACATCATGCTCTTCAATGTTGTGCGTGGAGAGATGATGATGATG 3720  
Db 8673 GACCTGAGTGTATTAACATCATGCTCTTCAATGTTGTGCGTGGAGAGATGATGATGATG 8732  
Qy 3721 AAAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3749  
Db 8733 AAAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8761

RESULT 14  
ABK91411  
ID ABK91411 standard; DNA; 9605 BP.  
XX  
XX  
XX ABK91411;



DT 15-NOV-2002 (first entry)  
XX Hepatitis C virus Con 1 isolate DNA.  
XX  
XX HCV, ds: Con 1, adaptive mutation; liver failure; cirrhosis;  
KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV,  
KM internal ribosome entry site; IRES; NS5A; HCV replication.  
XX  
XX Hepatitis C virus.  
OS  
FH Key Location/Qualifiers  
FT CDS 342..9374  
FT /tag= a  
FT /product= "HCV polyprotein"  
FT /note= "The polyprotein consists of the Core, E1, E2, P7,  
FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
XX  
PN WO200259321-A2.  
XX  
XX 01-AUG-2002.  
XX  
XX 16-JAN-2002; 2002WO-EP000526.  
XX  
XX 23-JAN-2001; 2001US-0263479P.  
XX  
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.  
XX  
XX De Francesco R, Migliaccio G, Paonessa G;  
XX  
XX WPI; 2002-599793/64.  
XX  
XX P-PSDB; ABG32451.  
XX  
XX  
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX  
XX  
XX Claim 9; Page 36-39; 69pp; English.  
XX  
XX The invention relates to nucleic acid molecules comprising altered HCV  
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
CC internal ribosome entry site (IRES) region coding for one or more NS3,  
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
CC are detailed in the specification. Also included are (1) an expression  
CC vector comprising a nucleotide sequence coding for the altered nucleic  
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
CC an HCV replicon enhanced cell or which containing a functional HCV replicon; (5)  
CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1,  
CC used as a basis for the adaptive mutations of the invention  
XX  
SQ Sequence 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 U; 0 Other;  
Query Match 87.8%; Score 3291.4; DB 6; Length 9605;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;  
QY 1 TGGAGGGGCGCTTACAGAGGCTCACCGACGTGGATCCCACTTCTGCTCCCAACAAG 60  
DB 5013 TGGAGAGGCGCTTTTACAGGCTTACCAATAGACGCCCAATTTCTGCTCCAGATCAAG 5072  
QY 61 CAGGACAGAGACAATTCCCTTACTGTTGGCGGTACAGGCTACTGTGTGCGGTAGGACC 120  
DB 5073 CAGGACAGAGACAATTCCCTTACTGTTAGCATACAGGCTACGGTGTGCGGCAAGGCT 5132

QY 121 CAGGCCCACTTCATATGAGATCAAAATGTGAAAGTGTCTCATACGGCTAAAGCTTACT 180  
DB 5133 CAGGCTCCACTTCATATGAGATCAAAATGTGAAAGTGTCTCATACGGCTAAAGCTTACG 5192  
QY 181 CTGCGCGGGCCCAACACCTTGTGTATAGGCTGGAGCCGCTCAAAAGAGGTACCCCTC 240  
DB 5193 CTGCAAGGGCCCAACGCCCTGTGTATAGGCTGGAGCCGCTCAAAAGAGGTATCTAC 5252  
QY 241 ACACACCCCATTAACAAATTCATATGAGATCATGTGACCGACCTGAGAGGTCTGTACG 300  
DB 5253 ACACACCCCATTAACAAATTCATATGAGATCATGTGACCGACCTGAGAGGTCTGTACG 5312  
QY 301 AGCACCTGGGTCTGGTGGGCGGGGTCTTGGACGCTTGGCTGCGTATTTGTTGACACA 360  
DB 5313 AGCACCTGGGTCTGGTGGGCGGGGTCTTGGACGCTTGGCTGCGTATTTGTTGACACA 5372  
QY 361 GGCAGCGTGTATTTGTGGGTAGATCATCTTGTCCGGCCGGCGCTATTGTTCCCGAC 420  
DB 5373 GGCAGCGTGTATTTGTGGGTAGATCATCTTGTCCGGAAACCGGCCATCATTTCCCGAC 5432  
QY 421 AGGGAAGTCTCTACAGAGATTTCATGATGAGAAAGATGGCGGTGACCTCTCTAC 480  
DB 5433 AGGGAAGTCTCTACAGAGATTTCATGATGAGAAAGATGGCGGTCTCTACCTCTTAC 5492  
QY 481 ATCGACAGGGAATGAGCTGCGCGAGCATTCAGCAAAAAGCGCTCGGTTGCTGACG 540  
DB 5493 ATCGAAGGGAATGAGCTGCGCGAGCAATTCAGCAAAAAGCGGTTGCTGCA 5552  
QY 541 ACAGCCCAACAGCAAGCGGAGCGCTGCTCCCGTGTGTGAGTCCAGTGGGAGCGCTT 600  
DB 5553 ACAGCCCAACAGCAAGCGGAGCGCTGCTCCCGTGTGTGAGTCCAGTGGGAGCGCTT 6612  
QY 601 GAGACCTTGGGGCGAAACAACATGTGGAATTCATCAGCGGGATACAGTACTTAGAGGCG 660  
DB 5613 GAGACCTTGGGGCGAAAGATGTGGAATTCATCAGCGGGATACAGTACTTAGAGGCG 5672  
QY 661 TTGTCACTCTGCTGGGAATCCCGGATGTGATCACTGATGAGCGCTTCAAGCGCTCTCTC 720  
DB 5673 TTGTCACTCTGCTGGGAATCCCGGATGTGATCACTGATGAGCGCTTCAAGCGCTCTAC 5732  
QY 721 ACTAGCGGCTACCAACCAATTCATCTCTGCTTTAATCTCTGGGGGATGGTATGCC 780  
DB 5733 ACTAGCGGCTACCAACCAATTCATCTCTGCTTTAATCTCTGGGGGATGGTATGCC 5792  
QY 781 GCCCAATGCTGCTCCCGGAGTGTGCTGATGAGTGTGAGCGGCGGATGTGTGAGCG 840  
DB 5793 GCCCAATGCTGCTCCCGGAGTGTGCTGATGAGTGTGAGCGGCGGATGTGTGAGCG 5852  
QY 841 GCTGTTGGAGATAGAGCTTGGGAAGTGTGTGATCATCTTGGCGGCTATAGAGCA 900  
DB 5853 GCTGTTGGAGATAGAGCTTGGGAAGTGTGTGATCATCTTGGCGGCTATAGAGCA 5912  
QY 901 GGAATGGAGGCGCGCTGTGGCTTTAAGTATAGAGCGGGAATGCTCTTCCACCGAG 960  
DB 5913 GGAATGGAGGCGCGCTGTGGCTTTAAGTATAGAGCGGGAATGCTCTTCCACCGAG 5972  
QY 961 GACTGTGTTAATCTTCTCTGCAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 5973 GACTGTGTTAATCTTCTCTGCAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 6032  
QY 1021 TGGCAGCGATACTGCTGGGCGGACGTGGGTCCAGGGGAGGGGCTTGGAGTGAAG 1080  
DB 6033 TGGCAGCGATACTGCTGGGCGGACGTGGGTCCAGGGGAGGGGCTTGGAGTGAAG 6092  
QY 1081 CGGCTATAGCGTTTCGCTTGGCGGGGTAAACATGTTTCCCAACGACATATGTGCGAG 1140  
DB 6093 CGGCTATAGCGTTTCGCTTGGCGGGGTAAACATGTTTCCCAACGACATATGTGCGAG 6152  
QY 1141 AGCGATCGCGAGACGTGTCACTGATATCTTCCGACCTTATATCAACCAATGTTG 1200  
DB 6153 AGCGATCGCGAGACGTGTCACTGATATCTTCCGACCTTATATCACTACTGAGCTGCTG 6212



QY 1201 AAGAGCTCCACGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1260  
DB 6213 AAGAGCTTCACGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6272  
QY 1261 AAGAGTGTGGGACGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1320  
DB 6273 AAGAGTGTGGGACGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6332  
QY 1321 AAGCTCTGCGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1380  
DB 6333 AAGCTCTGCGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6392  
QY 1381 GTCTGCGGAGGAGACGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1440  
DB 6393 GTCTGCGGAGGAGACGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6452  
QY 1441 CATGTCAAAAACGCTTCATGAGAGTGTGGGCTTAAGACCTGTAGTAAACATGTGGCAT 1500  
DB 6453 CATGTCAAAAACGCTTCATGAGAGTGTGGGCTTAAGACCTGTAGTAAACATGTGGCAT 6512  
QY 1501 GGAACATTCGCCCATCAACGATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1560  
DB 6513 GGAACATTCGCCCATTAACGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6572  
QY 1561 TATTCAGAGGCGCTGTGCGGAGTGTGGGCTTAAGACCTGTAGTAAACATGTGGCAT 1620  
DB 6573 TATTCAGAGGCGCTGTGCGGAGTGTGGGCTTAAGACCTGTAGTAAACATGTGGCAT 6632  
QY 1621 GATTTCACTACGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1680  
DB 6633 GATTTCACTACGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6692  
QY 1681 GCGCCGGAATTCCTCAACAGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1740  
DB 6693 GCGCCGGAATTCCTCAACAGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6752  
QY 1741 AAACCTCTCTCAACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1800  
DB 6753 AAACCTCTCTCAACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6812  
QY 1801 TGCGAGCTCCCATGAGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1860  
DB 6813 TGCGAGCTCCCATGAGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6872  
QY 1861 CCTCCCATCAACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1920  
DB 6873 CCTCCCATCAACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6932  
QY 1921 TTGGCAGCTCTCAACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 1980  
DB 6933 TTGGCAGCTCTCAACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 6992  
QY 1981 CAAATGAGCTTCACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2040  
DB 6993 CAAATGAGCTTCACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7052  
QY 2041 GCGCGGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2100  
DB 7053 GCGCGGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7112  
QY 2101 CCGCTCCAGGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2160  
DB 7113 CCGCTCCAGGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7172  
QY 2161 TCCAGAAATTCACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2220  
DB 7173 TCCAGAAATTCACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7232  
QY 2221 CTGAGATCTGAGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2280  
DB 7233 CTGAGATCTGAGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7292  
QY 2281 CCTACTAAGACCCCTCTATACCACTCAACGAGGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2340

DB 7293 CCTGCAAGGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7352  
QY 2341 TCCAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2400  
DB 7353 TCCAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7412  
QY 2401 TCCAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2460  
DB 7413 TCCAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7472  
QY 2461 GAGAGTCTGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2520  
DB 7473 GAGAGTCTGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7532  
QY 2521 CCGGATCTGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2580  
DB 7533 CCGGATCTGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7592  
QY 2581 TGCTGCTGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2640  
DB 7593 TGCTGCTGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7652  
QY 2641 AGCAAGTCCCATCAACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2700  
DB 7653 AGCAAGTCCCATCAACGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7712  
QY 2701 GCTACCAATCCCGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2760  
DB 7713 GCTACCAATCCCGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7772  
QY 2761 ATCTGAGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2820  
DB 7773 ATCTGAGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7832  
QY 2821 AAGGCTAAGCTTCTATCAAGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2880  
DB 7833 AAGGCTAAGCTTCTATCAAGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7892  
QY 2881 TCTAAATTTGGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 2940  
DB 7893 TCTAAATTTGGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 7952  
QY 2941 ATCCGCTCCGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 3000  
DB 7953 ATCCGCTCCGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 8012  
QY 3001 ATGGCAAAAATGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 3060  
DB 8013 ATGGCAAAAATGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 8072  
QY 3061 CTATGCTGTTCCAGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 3120  
DB 8073 CTATGCTGTTCCAGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 8132  
QY 3121 GTCTCAACCTCCAGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 3180  
DB 8133 GTCTCAACCTCCAGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 8192  
QY 3181 CAGCGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 3240  
DB 8193 CAGCGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 8252  
QY 3241 TATGACACCCGCTGTTTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 3300  
DB 8253 TATGACACCCGCTGTTTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 8312  
QY 3301 ATTATCAATGTTGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 3360  
DB 8313 ATTATCAATGTTGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 8372  
QY 3361 CCGCTTATATGAGGAGTGAATTAACGAGAGCTGCTCCAGCCCTTCGCGCTGCTGCTA 3420

Db 8373 CGGCTTTACATCGGGGGCCCCCTGACTAATTTTAAAGGCGAGAACTCGCGCTATCGCCGG 8432  
Qy 3421 TGGCCGCGAGAGGGCGCTGCTGACGACTAGCTGGATATACCTTCATGTTACTTGAAG 3480  
Db 8433 TGGCCGCGAGAGGGCTTACTGACGACGACTGCGGTAATACCTTCATGTTACTTGAAG 8492  
Qy 3481 GCGCTGACAGCCTTGAGCTGAGCTGCAAGCTTCAGAGACTGACGATGCTGTCGCGAGAC 3540  
Db 8493 GCGCGTGGGGCTGTGAGCTGGAAGCTTCAGAGCTGACGATGCTGTCGTCGATGCGAGAC 8552  
Qy 3541 GCGCTGTGCTTATCTGTGAGAGCGGCGGAACCCAGAGAGACCGCGGACCTTACGAGATC 3600  
Db 8553 GACCTGTGCTTATCTGTGAGAGCGGCGGAACCCAGAGAGACCGCGGACCTTACGAGATC 8612  
Qy 3601 TTCAGGAGGCTATGACTGATGCTGCTGCCCCCGGGGAGACCGCGCCCAACGAGATAC 3660  
Db 8613 TTCAGGAGGCTATGACTGATGCTGCTGCCCCCGGGGAGACCGCGCCCAACGAGATAC 8672  
Qy 3661 GACCTGAGATGATTAATCATCATGCTCTCCATGATGTCGCTCGCGACGATGATCTGGC 3720  
Db 8673 GACTTGAGATGATTAATCATCATGCTCTCCATGATGTCGCTCGCGACGATGATCTGGC 8732  
Qy 3721 AAAAGGATATCTACTCTACCTACCCGCTGACC 3749  
Db 8733 AAAAGGATATCTACTCTACCTACCCGCTGACC 8761

RESULT 15  
ID ABR91425 standard; DNA; 9605 BP.  
XX ABR91425;  
AC ABR91425;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE Hepatitis C virus Con 1 isolate DNA mutant 2.  
XX  
KM HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KM internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
XX  
OS Hepatitis C virus.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 342..9374  
FT FT /\*tag= a  
FT /product= "HCV polyprotein"  
FT /note= "The polyprotein consists of the Core, E1, E2, P7,  
FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
FT mutation replace(3946,A)  
FT /\*tag= b  
XX  
XX WO200259321-A2.  
XX PN  
XX PD 01-AUG-2002.  
XX  
PF 16-JAN-2002; 2002WO-EP000526.  
XX  
XX 23-JAN-2001; 2001US-0263479P.  
XX PR  
XX (RICE-) IST RICERCH BIOL MOLECOLARE ANGELUCCI.  
XX PA  
XX PI De Francesco R, Migliaccio G, Paonessa G;  
XX  
XX WPI; 2002-599793/64.  
XX  
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
XX expression.  
XX  
PS Claim 9; Page; 69pp; English.

XX  
CC The invention relates to nucleic acid molecules comprising altered HCV  
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
CC internal ribosome entry site (IRES) region coding for one or more NS3,  
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
CC are detailed in the specification. Also included are (1) an expression  
CC vector comprising a nucleotide sequence coding for the altered nucleic  
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC proteins, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
CC mutant of the invention. Note: The present sequence is not shown in the  
CC specification but was created by the indexer using the HCV sequence  
CC appearing as ABR91411 and the information in Claim 9  
XX  
SQ Sequence 9605 BP; 1909 A; 2883 C; 2734 G; 2079 T; 0 U; 0 Other;  
XX  
Query Match 87.8%; Score 3291.4; DB 6; Length 9605;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;  
Qy 1 TGGAGAGGCGGCTTTCACAGGCTTCACCCAGTGAATGCCACTTCTGTCCTCCAAACAG 60  
Db 5013 TGGAGAGGCGGCTTTCACAGGCTTCACCCAGTGAATGCCACTTCTGTCCTCCAAACAG 5072  
Qy 61 CAGGACGAGAGAACTTCCCTACCTGCTGGGCTGACCAAGCTACTGTGGCTGGGCG 120  
Db 5073 CAGGACGAGAGAACTTCCCTACCTGCTGGGCTGACCAAGCTACTGTGGCTGGGCG 5132  
Qy 121 CAGGCCCCACCTTCATCATGAGATCAATGTGAAGTGTCTATACGCTTAAAGCTTACT 180  
Db 5133 CAGGCTCCACCTTCATCATGAGATCAATGTGAAGTGTCTATACGCTTAAAGCTTAA 5192  
Qy 181 CTGGCGGGGCAACACCTTGTCTGTATAGGCTGGGAGCCGCTCAAAAAGAGTCAACCTTC 240  
Db 5193 CTGGACGGGCAACACCGCCCTGTGTATAGGCTGGGAGCCGCTCAAAAAGAGTCAAC 5252  
Qy 241 ACACACCCCATTAACCAATTCATCATGAGATCAATGTGAAGTGTCTTAAAGCTTAA 300  
Db 5253 ACACACCCCATTAACCAATTCATCATGAGATCAATGTGAAGTGTCTTAAAGCTTAA 5312  
Qy 301 AGCAGCTGGGTGCTGGTGGGCGGGGCTTGGAGCTCTGGCTGCGATTGTGACAA 360  
Db 5313 AGCAGCTGGGTGCTGGTGGGCGGGGCTTGGAGCTCTGGCTGCGATTGTGACAA 5372  
Qy 361 GGCAGGCTGATCATGTGGGTAGATCATCTTGTCCGGGCGCGGCTATTGTCCGCAC 420  
Db 5373 GGCAGGCTGATCATGTGGGTAGATCATCTTGTCCGGGCGCGGCTATTGTCCGCAC 5432  
Qy 421 AGGGAAGTCTCTTACAGAGATTCATGAGATGGAAGATGGCGCTCGACCTCCCTTAC 480  
Db 5433 AGGGAAGTCTCTTACAGAGATTCATGAGATGGAAGATGGCGCTCGACCTCCCTTAC 5492  
Qy 481 ATGAGACAGGGAATGAGCTCGCGAGAGTTCAGCAAAAGCGCTCGGGTTGTGACAG 540  
Db 5493 ATGAGACAGGGAATGAGCTCGCGAGAGTTCAGCAAAAGCGCTCGGGTTGTGACAG 5552  
Qy 541 ACAGCCACCAAGCAAGCAGAGCGCGCTGCTCCGCTGGTGAAGTCCAAATGGCGAGCCTT 600  
Db 5553 ACAGCCACCAAGCAAGCAGAGCGCGCTGCTCCGCTGGTGAAGTCCAAATGGCGAGCCTT 5612  
Qy 601 GAGACCTTCTGGCGGCAAAACATATGGAATTCATACAGCGGAGTACAGTACTTAAAGAG 660  
Db 5613 GAGACCTTCTGGCGGCAAGATATGGAATTCATACAGCGGAGTACAGTACTTAAAGAG 5672

QY	661	TTGTCCACTCTGCTCGGGAAATCCCGGATTGCATCTGATGCGCTTCAACAGCTCTGTG	720
Db	5673	TTGTCCACTCTGCTCGGGAAACCCCGCATGATCATGATGGCAATTCACAGCCTCTATC	5732
QY	721	ACTAGCCCGCTACCAACCAATCTACCCCTCGCTTAATCATCTCGGGGGAAATGGGTACC	780
Db	5733	ACGACCCGCTACCAACCAATACCTCTCTGTTAACTCTCGGGGAATGGGTGGC	5792
QY	781	GCCCACTGCTCCCCCAGTGTCTTCAAGCTTTCGTAGCGCGGCAATGCTGTGCG	840
Db	5793	GCCCAACTGCTCTCCAGCGCTCTTCTGTGTTTCGTAGCGCGGCAATGCTGTGCG	5852
QY	841	GCTGTGGCAGATATAGGCTTGGGAAAGTGTGTGACATCTTGGCGGCTATGAGCA	900
Db	5853	GCTGTGGCAGATATAGGCTTGGGAAAGTGTGTGAAATTTTGGCAGGTTATGAGCA	5912
QY	901	GGAGTGGCAGGCGCGCTGTGGCTTTTAAAGTCATGACCGGCGAATCCCTCCACGAG	960
Db	5913	GGGAGTGGCAGGCGCGCTGTGGCTTTTAAAGTCATGACCGGCGAATCCCTCCACGAG	5972
QY	961	GACCTGGTTAACTTACTCCCTGCATCTCTCTCTGTGCCCTGTGTGTGGGGTCTGT	1020
Db	5973	GACCTGGTTAACTTACTCCCTGTCTATCTCTCTCCCTGCGGCTTAAGTGTGGGGTCTGT	6032
QY	1021	TGCGCAGCGAATACGTGTCGCGCAGTGGGTCCAGGGGAGGGGGCTGTGACATGATGAC	1080
Db	6033	TGCGCAGCGAATACGTGTCGCGCAGTGGGTCCAGGGGAGGGGGCTGTGACATGATGAC	6092
QY	1081	CGGCTGATAGCGTTTGCCCTCGCGGGGTTAACATGTTTCCGCCACGACTATGTGCAAG	1140
Db	6093	CGGCTGATAGCGTTTGCCCTCGCGGGGTTAACACAGCTCTCCCGACGACTATGTGCTAG	6152
QY	1141	AGCGACGCGCAGACGCTGCTACCTAGATCCTCTCCGACCTTACTATACCCCACTGTG	1200
Db	6153	AGCGACGCTGACACGCTGCTACCTAGATCCTCTCTTAACTTAACTACTACTGCTGTG	6212
QY	1201	AAGAGGCTCAACCAATGATTAAACAGGACTGTCCAGGCCCTGCTCGGCTCGTGCTA	1260
Db	6213	AAGAGGCTTCAACCAATGATCAACAGGACTGTCTCAAGCCATGTCCGGCTCGTGCTA	6272
QY	1261	AGGATGTTTGGGACTGGAATATGACAGATTTTGGCTGACTTCAAGACTGTGCTCAAGTCC	1320
Db	6273	AGAGATGTTTGGGATTTGGATATGACAGGCTGTGACTGATTTCAAGACTGTGCTCAAGTCC	6332
QY	1321	AAGCTCCGCGCGCAATTCACGGGAGTCCCTTTTCTATGCAACGATGGGTTACAAGGG	1380
Db	6333	AAGCTCCGCGCGCAATTCACGGGAGTCCCTTTCTTCTATGTCACAGTGGGTTACAAGGA	6392
QY	1381	GTTGCGCGGGAGACGGGATCATGCAACCACTGTCTATGTGAGGACAGATCACCGGA	1440
Db	6393	GTTGCGCGGGAGACGGGATCATGCAACCACTGTCCATGTGAGGACAGATCACCGGA	6452
QY	1441	CATGTCAAAAACGTTTCCATAGAGATCTTGGGCTTAAGACTGTATGTAATATGTGCAT	1500
Db	6453	CATGTCAAAAACGTTTCCATAGAGATCTTGGGCTTAAAGACTGTATGTAACATGTGCAT	6512
QY	1501	GGAACATTTCCCATTTAACGCAATACCAACGAGGCCCTTCAGCCTCCCGACGCCAATC	1560
Db	6513	GGAACATTTCCCATTTAACGCGTAACCAACGAGGCCCTTCAGCCTCCCGCGGCCAAT	6572
QY	1561	TATTCACAGGCGCTGTGCGGCGTGTCTGTAGAGATACGTGAGAGTTTACGCGGCGTGGG	1620
Db	6573	TATTCACAGGCGCTGTGCGGCGTGTCTGTAGAGATACGTGAGAGTTTACCGCGGCGTGGG	6632
QY	1621	GATTTCCACTACGTGACAGAGCATGACCACTGACAAACGTAAAAATGCCGTGACAGTTCCA	1680
Db	6633	GATTTCCACTACGTGACAGGCGATGACCACTGACAAACGTAAAGTCCCGTGTCAAGTTCCG	6692
QY	1681	GCCCCCGAATTTCTTACAGAAATGAGTGGGGTGGCGGCTGCAACAGTACGCTCGGCGTGC	1740
Db	6693	GCCCCCGAATTTCTTACAGAAATGAGTGGGGTGGCGGCTGCAACAGTACGCTCGGCGTGC	6752

QY	1741	AAACCTCTCTCAAGGAGAGAGATGACATTTCCAGGTGGGCTCAACCAATACCTGTTGGG	1800
Db	6753	AAACCCCTCTCAACGGAGAGAGATGACATTTCTGGTGGGCTCAATCAATACCTGTTGGG	6812
QY	1801	TGCAGAGCTCCATGCGAGGCCGGAACCGGAAATGACAGAGCTCACTTCATGCTCACGGAC	1866
Db	6813	TCACAGCTCCATGCGAGGCCGGAACCGGAACTGACAGAGCTCACTTCATGCTCACGGAC	6872
QY	1861	CCCTCCCAATCAAGCAGAGACGGCTTAAGCGCAGAGCTGGCCAGGGGGTCTCCCCCTCC	1920
Db	6873	CCCTCCCAATTAAGCGCGAGACGGCTTAAGGTGAGGTGGCCAGGGGATCTCCCCCTCC	6932
QY	1921	TTGGCAGAGCTTTAGCTAGCTAGCCAGTGTCTGGCCTTCTCGAAGCGCATACATTACC	1980
Db	6933	TTGGCAGAGCTATAGCTAGCTAGCCAGTGTCTGGCCTTCTCGAAGCGCATACATTACC	6992
QY	1991	CAAAATGACTTCCAGAGCTGACCTCATGAGGCGCAACCTCGTGGCGGCAATGAAATG	2040
Db	6993	CGTCTGACTCCCCGAGAGCTGACCTCATGAGGCGCAACCTCGTGGCGGCAATGAAATG	7052
QY	2041	GGCGGGGACATTACCCCGCGTGGAGTTCAGAGAAACAAGTAGTAATCTTGACTTTTGAC	2100
Db	7053	GGCGGGAAATCAACCCCGGTGGAGTTCAGAAATAAGTAGTAATTTTGACTCTTTGAG	7112
QY	2101	CCGCTCCGAGCGGAGAGAGATGAGCGGGAAGTGTCCGTCCCGCGGAGATCTCGGAAA	2160
Db	7113	CCGCTCCAAAGCGGAGAGAGATGAGAGGGAAGTATCCGTTCCGGCGGAGATCTCGGAGG	7172
QY	2161	TCCAAGAAATTTCCCAACAGCGCATGCCCCGATGGGCAAGCCCGGATTACACCTTCGGTG	2220
Db	7173	TCCAAGAAATTTCCCTCGAGCATATCCCATATGAGGCAAGCCCGGATTACACCTTCGACTG	7232
QY	2221	CTGAGTCTCTGGAAGCGCCCGGACATAGTCCCTCAGTGGTACATGAGGTGCCACTGCA	2280
Db	7233	TTAAGTCTCTGGAAGAGACCGGACATAGTCCCTCAGTGGTACAGCGGTCTCCATTGGC	7292
QY	2281	CCTACTAAGACCCCTCTTATACCACTTCCACGAGAGAAAGACAGTTGTTCTGACAGAA	2340
Db	7293	CCTGCAAGGCGCCCTCGATACCACTTCCACGAGAGAAAGAGCGTTGTTCTGTCAGAA	7352
QY	2341	TCACACGTGTCTTCTGCCCTGCGGAGCTTGCCACAAAGGCTTTCGTAAGTCCGAACCG	2400
Db	7353	TCATACGTGTCTTCTGCCCTTGGCGGACCTGCGCAAAAGACCTTCGCGAGTCCGAATCG	7412
QY	2401	TGGGCGGTGCAAGGGGCAAGGGCAACGGCCCTCTGACCAACCTCTCGACAGAGGGGA	2460
Db	7413	TCGGCGGTGCAAGGGGCAAGGGCAAGGGCCCTCTGACCAACCTCTCGACAGAGGGGAC	7472
QY	2461	GCAGAGTCTGAGCTTGAAGTCGTATCTCTCATGCCCCCTTGAGGGGGAGCCGGGGAC	2520
Db	7473	GCGGGATTCGACGTTGAAGTCGTATCTCTCATGCCCCCTTGAGGGGGAGCCGGGGAT	7532
QY	2521	CCCGATCTCAGCGACGGGTCCTTGCTTACCGGTGATGAGAGCGGTGAGAGCTGTC	2580
Db	7533	CCCGATCTCAGCGACGGGTCCTTGCTTACCGGTATGACGATGAGAGCGGTGAGAGCTGTC	7592
QY	2581	TGCGTCTGAGTCTTACATGAGCAGGGGCTCTGATTCACGCATGCGTCGGGAGAA	2640
Db	7593	TGCGTCTGAGTCTTACATGAGCAGGGGCTCTGATTCACGCATGCGTCGGGAGAA	7652
QY	2641	AGCAAGCTGCCATCAACGCGTTGAGCACTTTTGTGCTGCTACCAACATGATGTATC	2700
Db	7653	ACCAAGCTGCCATCAATGACACTGAGGAATCTTTTGTGCTGCTACCAACATGATGTATC	7712
QY	2701	GCTACCAATTCGGAGCGCAAGCCAGCGGCAAGAAAGTCACTTTGACAGCTGCA	2760
Db	7713	GCTACCAATTCGCGCAAGCGCAAGCTCGGCAAGAAAGTCACTTTGACAGCTGCA	7772
QY	2761	ATCTGTGACATTCATCAACAGACCTGTCTAAGAGATGAAAGCGAAGGGTCCACAGTT	2820
Db	7773	GTCCTGTGACATTCATCAACGGAGAGTGTCTAAGAGATGAAAGGGTCCACAGTT	7832
QY	2821	AAGCTTAAGCTTCTATAGTAGAGAAAGCTGCAAGTGAAGCCCCCACTTGGCCAA	2880

```
Db 7833 AAGGCTAAACTTCTATCCGTGAGGAAAGCTGTAGTGACGCCCCACATTCGGCCAG 7892
QY 2881 TCTAAATTTGGCTATGGGGGAAAGAGCTCCGGAACCTATCCAGCAAGGCCATTAAACAC 2940
Db 7893 TCTAAATTTGGCTATGGGGGAAAGAGCTCCGGAACCTATCCAGCAAGGCCATTAAACAC 7952
QY 2941 ATCCGCTCCGTGTGGAGGAGCTGTGTGAAGACACTGAACACCAATTGACACCAACATC 3000
Db 7953 ATCCGCTCCGTGTGGAGGAGCTGTGTGAAGACACTGAAGACCAATTGACACCAACATC 8012
QY 3001 ATGGCAAAAAATGAGGTTTTCTGCGTCAAACGAGAGAGAGGCCGCAAGCCAGCTCCG 3060
Db 8013 ATGGCAAAAAATGAGGTTTTCTGCGTCAAACGAGAGAGAGGCCGCAAGCCAGCTCCG 8072
QY 3061 CTTATGTGTCTCCCAACTGGGGGTCGGTGTGTGCGAAGAAATGGCCCTTATGACGTG 3120
Db 8073 CTTATGTATTTCCCAATTTGGGGGTTCTGTGTGTGCGAAGAAATGGCCCTTATGACGTG 8132
QY 3121 GTCTCAACCTCCCTCAAGGCTGTGATGGGCTCCCTGATCGGATTCAGTATTCCTCTGA 3180
Db 8133 GTCTCAACCTCCCTCAAGGCTGTGATGGGCTCTTCATACGGAATTCATATCTCTCTGA 8192
QY 3181 CAGCGGGTCGAGTTCTGTGTGAACGCGTGGAAATCAAAAGAGACCCCTATGGCTTTGCA 3240
Db 8193 CAGCGGGTCGAGTTCTGTGTGAATGCTGTGAAGCGAAGAAATGCCCTATGGGCTTCGCA 8252
QY 3241 TATGACACCCGCTGTGTTTGAAGTCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 3300
Db 8253 TATGACACCCGCTGTGTTTGAAGTCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 8312
QY 3301 ATTATCAATGTGTGACTTGGGCCCCCGAAGCCAGACAGGCCATTAAGTGCCTCACAGAG 3360
Db 8313 ATCTACCAATGTGTGACTTGGGCCCCCGAAGCCAGACAGGCCATTAAGTGCCTCACAGAG 8372
QY 3361 CGGCTTTATATCGGGGGTCCCTGACTAATTCAAAAGGCGAGAACTGCGGCTATCGCCGG 3420
Db 8373 CGGCTTTATATCGGGGGTCCCTGACTAATTCTAAAGGCGAGAACTGCGGCTATCGCCGG 8432
QY 3421 TGCCTGCGGAGCGCGGCTGTGAGAGTCAAGTCACTGCGGTATATCCCTCACATGTACTTGAAG 3480
Db 8433 TGCCTGCGGAGCGCGGCTGTGAGAGTCAAGTCACTGCGGTATATCCCTCACATGTACTTGAAG 8492
QY 3481 GCGCTGTGCAAGCTGTGAGGCTGCAAGGCTCCAGGACTGCAAGTGTGCTGTGTGCGAGAG 3540
Db 8493 GCGCTGTGCAAGCTGTGAGGCTGCAAGGCTCCAGGACTGCAAGTGTGCTGTGTGCGAGAG 8552
QY 3541 GGCCTTGTCTTATCTGTGAGAGCGCGGAAACCCAGAGAGACGCGGAGCTTACGAGTC 3600
Db 8553 GACCTTGTCTTATCTGTGAAAGCGCGGAGACCCAGAGAGACGAGGAGCTTACGAGGCC 8612
QY 3601 TTCAAGGAGGCTATGACTAGTACTGTGCCCCCCCCCGGGGACCCGCCCAACCAAGATAC 3660
Db 8613 TTCAAGGAGGCTATGACTAGTACTGTGCCCCCCCCCGGGGACCCGCCCAACCAAGATAC 8672
QY 3661 GACCTGAGTTGATTAATCATATGCTCTCCATGTGTGCGTGTGCGGACGATGCACTGTGCG 3720
Db 8673 GACCTGAGTTGATTAATCATATGCTCTCCATGTGTGCAATGTGTCAATCGCGACGATGCACTGTGCG 8732
QY 3721 AAAAGGATATCTACTACCCGTGACCC 3749
Db 8733 AAAAGGATATCTACTACCCGTGACCC 8761
```

Search completed: February 21, 2005, 11:17:06  
Job time : 1872 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 10:33:33 ; Search time 584 Seconds  
(without alignments)  
10506.914 Million cell updates/sec

Title: US-09-664-363-22

Perfect score: 3750  
Sequence: 1 TGGAGGGGGCTTTCACAGG.....ACTACTCACCCTGACCG 3750

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	3750	100.0	3750	3 US-08-191-160-22	Sequence 22, Appl
2	3292.2	87.8	8642	4 US-10-029-907-2	Sequence 2, Appl
3	3291.4	87.8	7989	4 US-09-539-601-10	Sequence 10, Appl
4	3291.4	87.8	8001	4 US-09-539-601-7	Sequence 7, Appl
5	3291.4	87.8	8637	4 US-09-539-601-4	Sequence 4, Appl
6	3291.4	87.8	8638	4 US-10-029-907-24	Sequence 24, Appl
7	3291.4	87.8	8639	4 US-10-029-907-1	Sequence 1, Appl
8	3291.4	87.8	8649	4 US-09-539-601-13	Sequence 13, Appl
9	3291.4	87.8	11076	4 US-09-539-601-1	Sequence 1, Appl
10	3288.2	87.7	8638	4 US-10-029-907-6	Sequence 6, Appl
11	3288.2	87.7	8638	4 US-10-029-907-7	Sequence 7, Appl
12	3288.2	87.7	8638	4 US-10-029-907-25	Sequence 25, Appl
13	3288.2	87.7	8638	4 US-10-029-907-5	Sequence 5, Appl
14	3285	87.5	8648	4 US-09-539-601-12	Sequence 22, Appl
15	3281.8	87.5	8001	4 US-09-539-601-25	Sequence 25, Appl
16	3281.8	87.5	11076	4 US-09-539-601-16	Sequence 16, Appl
17	3280.2	87.5	8001	4 US-09-539-601-19	Sequence 19, Appl
18	3280.2	87.5	11076	4 US-08-324-977-35	Sequence 35, Appl
19	3278.6	87.4	7863	1 US-08-324-977-35	Sequence 35, Appl
20	3278.6	87.4	7863	1 US-08-324-977-35	Sequence 35, Appl
21	3278.6	87.4	7863	2 US-08-324-977-35	Sequence 35, Appl
22	3278.6	87.4	7863	3 US-09-315-850-35	Sequence 35, Appl
23	3278.6	87.4	7917	1 US-08-324-977-31	Sequence 31, Appl
24	3278.6	87.4	7917	2 US-08-384-616-31	Sequence 31, Appl
25	3278.6	87.4	7917	2 US-08-304-686A-31	Sequence 31, Appl
26	3278.6	87.4	7917	3 US-09-315-850-31	Sequence 31, Appl
27	3278.6	87.4	9030	1 US-08-324-977-13	Sequence 13, Appl

28	3278.6	87.4	9030	2 US-08-384-616-13	Sequence 13, Appl
29	3278.6	87.4	9030	2 US-08-304-686A-13	Sequence 13, Appl
30	3278.6	87.4	9030	3 US-09-315-850-13	Sequence 13, Appl
31	3278.6	87.4	9416	1 US-08-324-977-1	Sequence 1, Appl
32	3278.6	87.4	9416	2 US-08-384-616-1	Sequence 1, Appl
33	3278.6	87.4	9416	2 US-08-304-686A-1	Sequence 1, Appl
34	3278.6	87.4	9416	3 US-09-315-850-1	Sequence 1, Appl
35	3278.6	87.4	9416	3 US-08-823-895A-27	Sequence 27, Appl
36	3275.4	87.3	8001	4 US-09-539-601-28	Sequence 28, Appl
37	3275.4	87.3	11076	4 US-09-539-601-31	Sequence 31, Appl
38	3230.6	86.1	9595	3 US-09-014-416-4	Sequence 4, Appl
39	3200.2	85.3	9413	4 US-09-827-688-6	Sequence 6, Appl
40	3192.2	85.1	9412	4 US-08-150-204E-96	Sequence 96, Appl
41	2449.6	65.3	12980	3 US-08-811-566-5	Sequence 5, Appl
42	2449.6	65.3	12980	3 US-09-034-756-5	Sequence 5, Appl
43	2448	65.3	9646	3 US-08-811-566-1	Sequence 1, Appl
44	2448	65.3	9646	3 US-09-034-756-1	Sequence 1, Appl
45	2446.4	65.2	9365	4 US-09-827-688-7	Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-08-191-160-22  
; Sequence 22, Application US/08191160  
; Patent No. 6210675  
; GENERAL INFORMATION:  
; APPLICANT: Highfield, Peter Edmund  
; APPLICANT: Rodgers, Brian Colin  
; APPLICANT: Tedder, Richard Seton  
; APPLICANT: Barbara, John Anthony James  
; TITLE OF INVENTION: Viral Agent  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSER: Rothwell, Figg, Ernst & Kurz  
; STREET: 1700 K Street  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage  
; OPERATING SYSTEM: IBM AT compatible  
; SOFTWARE: Wordperfect 5.0 (DOS text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/191,160  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/628,516  
; FILING DATE: 17 DEC 1990  
; APPLICATION NUMBER: UK 89 28 562.1  
; FILING DATE: 18 DEC 1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 414.0  
; FILING DATE: 27 FEB 1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 814.1  
; FILING DATE: 03 MAR 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: E. Anthony Figg  
; REGISTRATION NUMBER: 27,195  
; REFERENCE/DOCKET NUMBER: 1645-103A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 833-5740  
; TELEFAX: (202) 833-5744  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3750 base pairs  
; TYPE: nucleotide with corresponding protein  
; STRANDEDNESS: single

/ TOPOLOGY: linear  
 / MOLECULE TYPE: cDNA to genomic RNA  
 / ORIGINAL SOURCE:  
 / ORGANISM: human; serum infectious for PT-NANBH  
 / IMMEDIATE SOURCE:  
 / LIBRARY: cDNA clones from 3' end of the genome  
 / FEATURE:  
 / LOCATION: from 1 to 3750 bp portion of the PT-NANBH  
 / LOCATION: polypeptide  
 / OTHER INFORMATION: viral non-structural proteins  
 / US-08-191-160-22

Query Match 100.0%; Score 3750; DB 3; Length 3750;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGAGGGCGCTTTACAGGCTTACCCAGTGTGATGCCACTTCTGTCCCAAAAG 60  
 DB 1 TGGGAGGGCGCTTTACAGGCTTACCCAGTGTGATGCCACTTCTGTCCCAAAAG 60  
 QY 61 CAGGAGGAGACAACTTCCCTACTGTGGCGGACAGGCTACTGTGTGCGTAGGGCC 120  
 DB 61 CAGGAGGAGACAACTTCCCTACTGTGGCGGACAGGCTACTGTGTGCGTAGGGCC 120  
 QY 121 CAGGCGCCACTTCATCATGTGATCAATGTGGAAGTGTCTATACGGCTTAAAGCTTACT 180  
 DB 121 CAGGCGCCACTTCATCATGTGATCAATGTGGAAGTGTCTATACGGCTTAAAGCTTACT 180  
 QY 181 CTGGCGGGGCGCAACCCCTTGTGTATAGCTGGAGCGCTCCAAAAGAGGTACCCCTC 240  
 DB 181 CTGGCGGGGCGCAACCCCTTGTGTATAGCTGGAGCGCTCCAAAAGAGGTACCCCTC 240  
 QY 241 ACAACACCCATTAACCAATTCATCATGTGATCAATGTGGAAGTGTCTATACGGCTTAAAGCTTACT 300  
 DB 241 ACAACACCCATTAACCAATTCATCATGTGATCAATGTGGAAGTGTCTATACGGCTTAAAGCTTACT 300  
 QY 301 AGGACCTGGTGTGTGGCGGGGCTCTTGAAGCTTGTGGTGTGATTTGCTTGAACA 360  
 DB 301 AGGACCTGGTGTGTGGCGGGGCTCTTGAAGCTTGTGGTGTGATTTGCTTGAACA 360  
 QY 361 GGCAGCGTGTGATTTGGTGGTGAATATCTTGTCCGGGCGGCGGCTTATTTGCCGAC 420  
 DB 361 GGCAGCGTGTGATTTGGTGGTGAATATCTTGTCCGGGCGGCGGCTTATTTGCCGAC 420  
 QY 421 AGGGAATCTCTTACCGAGAGTTCGATGAGTGAAGTGAAGTGGCGGTGCACTCCCTTAC 480  
 DB 421 AGGGAATCTCTTACCGAGAGTTCGATGAGTGAAGTGAAGTGGCGGTGCACTCCCTTAC 480  
 QY 481 ATCGAGGAGGAAATGACAGCTCGCCGAGCAAGTTCAGCAAAAAGCGCTCGGGTTGCTGAC 540  
 DB 481 ATCGAGGAGGAAATGACAGCTCGCCGAGCAAGTTCAGCAAAAAGCGCTCGGGTTGCTGAC 540  
 QY 541 ACAGCCACCAAGAGCGGAGCGGCTGCTCCGCTGTGATGATCAAGTGGCGAGCCCTT 600  
 DB 541 ACAGCCACCAAGAGCGGAGCGGCTGCTCCGCTGTGATGATCAAGTGGCGAGCCCTT 600  
 QY 601 GAGACCTTGGGGGAAACATGATGGAACCTTCAACGCGGGAATACAGTACTTGAAGAGG 660  
 DB 601 GAGACCTTGGGGGAAACATGATGGAACCTTCAACGCGGGAATACAGTACTTGAAGAGG 660  
 QY 661 TTGTCACACTGCTGGGAAATCCGCGATTTGCACTGATGAGGCTTCAAGGCTCTGTG 720  
 DB 661 TTGTCACACTGCTGGGAAATCCGCGATTTGCACTGATGAGGCTTCAAGGCTCTGTG 720  
 QY 721 ACTAGCCGCTCACCAACCAATCTTAACTCTGCTTAACTCTGCGGAGGATGGGTAGCC 780  
 DB 721 ACTAGCCGCTCACCAACCAATCTTAACTCTGCTTAACTCTGCGGAGGATGGGTAGCC 780  
 QY 781 GCCCAATCTGCTCCCGGAGTGTGCTTACGCTTCCGAGGCGCGGCAATGCTGTGCG 840  
 DB 781 GCCCAATCTGCTCCCGGAGTGTGCTTACGCTTCCGAGGCGCGGCAATGCTGTGCG 840  
 QY 841 GCTGTGGCAGCATAGGCTTGGGAAGGTGCTGTGACATCTTGGCGGCTATGAGCA 900

DB 841 GCTGTGGCAGCATAGGCTTGGGAAGGTGCTGTGACATCTTGGCGGCTATGAGCA 900  
 QY 901 GGAAGTGGAGGCGCGCTGCTGAGCTTTTAAAGTCATAGAGCGGAATGCCCTCCACCGAG 960  
 DB 901 GGAAGTGGAGGCGCGCTGCTGAGCTTTTAAAGTCATAGAGCGGAATGCCCTCCACCGAG 960  
 QY 961 GACTGTGTTAACTTACTCCCTGCACTCTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
 DB 961 GACTGTGTTAACTTACTCCCTGCACTCTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
 QY 1021 TGGCAGCGCATCTGCTGTGGGATGAGTGGGATCCAGGGAGGGGCTGTGAGTGAAG 1080  
 DB 1021 TGGCAGCGCATCTGCTGTGGGATGAGTGGGATCCAGGGAGGGGCTGTGAGTGAAG 1080  
 QY 1081 CGGCTGATGAGCTTGGCTCGCGGGGATACATGTTTCCCGCAGCAGCATATGTGCAGAG 1140  
 DB 1081 CGGCTGATGAGCTTGGCTCGCGGGGATACATGTTTCCCGCAGCAGCATATGTGCAGAG 1140  
 QY 1141 AGCGAGCGCGCAGCAGTGTCACTAGATCTCTCGACCTTATATCAACCACTGTTG 1200  
 DB 1141 AGCGAGCGCGCAGCAGTGTCACTAGATCTCTCGACCTTATATCAACCACTGTTG 1200  
 QY 1201 AAGAGCTCACCAATGATTAACAGAGACTGCTCAAGCCCTGCTCGGGCTGTGGCTTA 1260  
 DB 1201 AAGAGCTCACCAATGATTAACAGAGACTGCTCAAGCCCTGCTCGGGCTGTGGCTTA 1260  
 QY 1261 AAGAGCTCACCAATGATTAACAGAGACTGCTCAAGCCCTGCTCGGGCTGTGGCTTA 1320  
 DB 1261 AAGAGCTCACCAATGATTAACAGAGACTGCTCAAGCCCTGCTCGGGCTGTGGCTTA 1320  
 QY 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCAAGTGGTGAACAAGGG 1380  
 DB 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCAAGTGGTGAACAAGGG 1380  
 QY 1381 GTCGTGGGGGAGACGGCATATGACAGACCACTGCTCATGTGAGACAGATCAACCGGA 1440  
 DB 1381 GTCGTGGGGGAGACGGCATATGACAGACCACTGCTCATGTGAGACAGATCAACCGGA 1440  
 QY 1441 CATGTCAAAAAGGTTCCATGAGGATGTTGGGCTTAAGACCTGTGATTAAGTGGCAT 1500  
 DB 1441 CATGTCAAAAAGGTTCCATGAGGATGTTGGGCTTAAGACCTGTGATTAAGTGGCAT 1500  
 QY 1501 GGAACATTTCCCATCAACCATCAACCAAGGCGCTGACGCGCTCCAGGCGCAAC 1560  
 DB 1501 GGAACATTTCCCATCAACCATCAACCAAGGCGCTGACGCGCTCCAGGCGCAAC 1560  
 QY 1561 TATTCAGGGCGCTGTGGGGGTGCTGCTGAGAGTATAGTGAAGTTACGGGGTGGGG 1620  
 DB 1561 TATTCAGGGCGCTGTGGGGGTGCTGCTGAGAGTATAGTGAAGTTACGGGGTGGGG 1620  
 QY 1621 GATTTCCATGATGAGGATGACACGACCAAGTAAATGCGCGTGGCAGGTTGCA 1680  
 DB 1621 GATTTCCATGATGAGGATGACACGACCAAGTAAATGCGCGTGGCAGGTTGCA 1680  
 QY 1681 GCCCGGAAATCTTCAAGAAAGTGAAGTGGGCTGCGCTCAAGTATCCTCGGGTGC 1740  
 DB 1681 GCCCGGAAATCTTCAAGAAAGTGAAGTGGGCTGCGCTCAAGTATCCTCGGGTGC 1740  
 QY 1741 AAACTCTCTCAAGGAGAGTCAATTCAGGTGGGCTCAACCAATTAAGTGGTGG 1800  
 DB 1741 AAACTCTCTCAAGGAGAGTCAATTCAGGTGGGCTCAACCAATTAAGTGGTGG 1800  
 QY 1801 TGGCAGTCCCAAGGAGCGGAGCGGATGATGAGTGTGCTCAATGCTCAACGAG 1860  
 DB 1801 TGGCAGTCCCAAGGAGCGGAGCGGATGATGAGTGTGCTCAATGCTCAACGAG 1860  
 QY 1861 CCTCCACATCAACAGAGAGAGGCTTAAGCGAGGCTGCGCAGGGGGTCTCCCTCC 1920  
 DB 1861 CCTCCACATCAACAGAGAGAGGCTTAAGCGAGGCTGCGCAGGGGGTCTCCCTCC 1920  
 QY 1921 TTGGCAGCTTTCACTAGCAAGTGTGCTGCGCTTCTCGAAGCGACATCATTTACC 1980

1921 TTGGCAGCTCTTACAGTACGACGAGTGTCTGCGCCTTCTCGAAGCGCACTATACC 1980  
1981 CAAATGACTTCCAGACGCTGACCTATGAGGCGCAACTCTGTGGCGGCAATGAGT 2040  
1981 CAAATGACTTCCAGACGCTGACCTATGAGGCGCAACTCTGTGGCGGCAATGAGT 2040  
2041 GGGGGGAGATTAACCGCGTGAAGTACAGAGAACAGGTAGTAACTCTGACTTTTCAG 2100  
2041 GGGGGGAGATTAACCGCGTGAAGTACAGAGAACAGGTAGTAACTCTGACTTTTCAG 2100  
2101 CCGCTCCGAGCGGAGAGATGAGCGGAAAGTGTCCGTCCGCGGAGATCTGCGGAAA 2160  
2101 CCGCTCCGAGCGGAGAGATGAGCGGAAAGTGTCCGTCCGCGGAGATCTGCGGAAA 2160  
2161 TCCAAAGAAATCCACAGAGATGCGCGCATGGGCAACGCGGAAATTAACAACCTCCGCTG 2220  
2161 TCCAAAGAAATCCACAGAGATGCGCGCATGGGCAACGCGGAAATTAACAACCTCCGCTG 2220  
2221 CTGAGTCTGGAAGGCGCGGACTAGTCCCTCCAGTGTGATGAGGTTCCCACTGCCA 2280  
2221 CTGAGTCTGGAAGGCGCGGACTAGTCCCTCCAGTGTGATGAGGTTCCCACTGCCA 2280  
2281 CCTACTAAGACCCCTCTATATACCACTCCAGAGAGAGAGACAGTTGTTCTGACAGAA 2340  
2281 CCTACTAAGACCCCTCTATATACCACTCCAGAGAGAGAGACAGTTGTTCTGACAGAA 2340  
2341 TCCACCGTGTCTTCCCTGCGGAGCTTGGCCACAAAGGCTTGGTGTAGTCCGAAACG 2400  
2341 TCCACCGTGTCTTCCCTGCGGAGCTTGGCCACAAAGGCTTGGTGTAGTCCGAAACG 2400  
2401 TCGGCGGTGAGACGCGGAGAGCGGCAACGCGCTCTGACCAACCTCCGAGAGAGAG 2460  
2401 TCGGCGGTGAGACGCGGAGAGCGGCAACGCGCTCTGACCAACCTCCGAGAGAGAG 2460  
2461 GCAGAGTCTGACGCTTGAATGTTCTTCATGCCCCCTTGAAGGAGAGCGGGGAGAC 2520  
2461 GCAGAGTCTGACGCTTGAATGTTCTTCATGCCCCCTTGAAGGAGAGCGGGGAGAC 2520  
2521 CCCGATCTCAGGACGCGGCTTGTGTCTACCGTGAAGAGAGAGCGGCTGAGAGAG 2580  
2521 CCCGATCTCAGGACGCGGCTTGTGTCTACCGTGAAGAGAGAGCGGCTGAGAGAG 2580  
2581 TGTGTCTGATGCTTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
2581 TGTGTCTGATGCTTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
2641 AGCAAGCTGCCATCAACGCGTGAAGCAACTTTTGTGTGCTGCTACCAACATGCTAC 2700  
2641 AGCAAGCTGCCATCAACGCGTGAAGCAACTTTTGTGTGCTGCTACCAACATGCTAC 2700  
2701 GCTACCAATCCCGACGCGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
2701 GCTACCAATCCCGACGCGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
2761 ATCTGAGAGATCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
2761 ATCTGAGAGATCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
2821 AAGGCTAAGCTTCTATCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
2821 AAGGCTAAGCTTCTATCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
2881 TCTAATTTGGCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
2881 TCTAATTTGGCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
2941 ATCCGCTCCGTGAG 3000  
2941 ATCCGCTCCGTGAG 3000  
2941 ATCCGCTCCGTGAG 3000  
3001 ATGGCAAAATGAGGTTTTTCTGCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
3001 ATGGCAAAATGAGGTTTTTCTGCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 3060

3061 CTATCGTGTCCAGACTTGGGGGTCCGCTGTGCGAGAGAAATGAGCCCTATGAGCTG 3120  
3061 CTATCGTGTCCAGACTTGGGGGTCCGCTGTGCGAGAGAAATGAGCCCTATGAGCTG 3120  
3121 GTCTCACCCTCCCTGAGGCTGTGATGAGGCTCTCTGACGATTCAGTATTCCTGGA 3180  
3121 GTCTCACCCTCCCTGAGGCTGTGATGAGGCTCTCTGACGATTCAGTATTCCTGGA 3180  
3181 GAGCGGTGAGTCTCTGTGAAACGCTGGAATCAAAAGAGAGAGAGAGAGAGAGAGAG 3240  
3181 GAGCGGTGAGTCTCTGTGAAACGCTGGAATCAAAAGAGAGAGAGAGAGAGAGAGAG 3240  
3241 TATGACACCGGCTTTTGAAGTCAAGTCACTGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
3241 TATGACACCGGCTTTTGAAGTCAAGTCACTGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
3301 ATTATCAATGTTGTGATCTTGGCCCGAAGCAGACAGGCAATAGTCTGACAGAG 3360  
3301 ATTATCAATGTTGTGATCTTGGCCCGAAGCAGACAGGCAATAGTCTGACAGAG 3360  
3361 CGGCTTATATGCGGGGTCCCTGACTAATTCAAAAGGAGAGAGAGAGAGAGAGAGAG 3420  
3361 CGGCTTATATGCGGGGTCCCTGACTAATTCAAAAGGAGAGAGAGAGAGAGAGAGAG 3420  
3421 TGCAGGAG 3480  
3421 TGCAGGAG 3480  
3481 GCCTGTGACGCTGTGAGCTGCAAGAGTCAAGAGTGCAGAGTGTGCTGTGCGAGAG 3540  
3481 GCCTGTGACGCTGTGAGCTGCAAGAGTCAAGAGTGCAGAGTGTGCTGTGCGAGAG 3540  
3541 GGCCTTGTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
3541 GGCCTTGTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
3601 TTCAGGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660  
3601 TTCAGGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660  
3661 GACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3720  
3661 GACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3720  
3721 AAAAGGATTAATCACTCAACCGTGAACCG 3750  
3721 AAAAGGATTAATCACTCAACCGTGAACCG 3750

RESULT 2  
US-10-029-907-2  
; Sequence 2, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 8642  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802)...(8407)  
; NAME/KEY: Variation



LOCATION: 6268  
OTHER INFORMATION: r = a or g  
NAME/KEY: variation  
LOCATION: 4446  
OTHER INFORMATION: r = a or g  
US-10-029-907-2

Query Match 87.8%; Score 3292.2; DB 4; Length 8642;  
Best Local Similarity 92.3%; Pred. No. 0;  
Matches 3462; Conservative 2; Mismatches 285; Indels 0; Gaps 0;

QY 1 TGGGAGGGGCTTTCACAGGCTTCACCCAGTGATGCCACTTCTGTCCTCAACAAAG 60  
DB TGGGAGGGGCTTTCACAGGCTTCACCCAGTGATGCCACTTCTGTCCTCAACAAAG 4105  
QY 61 CAGGCGAGGAGCAACTTCCCTACCTGCTGAGGCTACAGGCTACTGCTGCTGAGGCT 120  
DB CAGGCGAGGAGCAACTTCCCTACCTGCTGAGGCTACAGGCTACTGCTGCTGAGGCT 4165  
QY 121 CAGGCGGCACTTCATCATGAGATCAAAATGTGGAAGTGTCTCATAGCGCTAAAGCTACT 180  
DB CAGGCGGCACTTCATCATGAGATCAAAATGTGGAAGTGTCTCATAGCGCTAAAGCTACT 4225  
QY 181 CTGCGCGGCGCAACACCTTGTGTATAGGCTGAGGCGCTCCAAAACGAGGTCACTC 240  
DB CTGCGCGGCGCAACACCTTGTGTATAGGCTGAGGCGCTCCAAAACGAGGTCTACTAC 4285  
QY 241 ACACACCCCAATCAATTCATCATGAGATGATGATGATGATGATGATGATGATGATG 300  
DB ACACACCCCAATCAATTCATCATGAGATGATGATGATGATGATGATGATGATGATG 4345  
QY 301 AGCAGCTGGGT 360  
DB AGCAGCTGGGT 4346  
QY 361 GGCAGCGTGTCTATTGT 420  
DB GGCAGCGTGTCTATTGT 4406  
QY 421 AGGGAAGTCTCTACCAAGAGGTTGATGATGATGATGATGATGATGATGATGATG 480  
DB AGGGAAGTCTCTACCAAGAGGTTGATGATGATGATGATGATGATGATGATGATG 4466  
QY 481 ATGAGAGGAGGAATGAGCTGCGGAGCACTTCAAGCAAAAGCGCTCGGTTGTCTGAG 540  
DB ATGAGAGGAGGAATGAGCTGCGGAGCACTTCAAGCAAAAGCGCTCGGTTGTCTGAG 4526  
QY 541 ACAAGCCACAAGCAAGCGAGGCGGCTGCTCCGCTGTGTGTGTGTGTGTGTGTGTGT 600  
DB ACAAGCCACAAGCAAGCGAGGCGGCTGCTCCGCTGTGTGTGTGTGTGTGTGTGTGT 4586  
QY 601 GAAGCCTTCTGGGGGAACAATGTGGAATTCATCAACGCGGATACATGATGATGATG 660  
DB GAAGCCTTCTGGGGGAACAATGTGGAATTCATCAACGCGGATACATGATGATGATG 4646  
QY 661 TTGTTCACCTGCTGCGGGAATCCCGAGATTGATCATCTGATGAGGTTCAAGCTCTGTC 720  
DB TTGTTCACCTGCTGCGGGAATCCCGAGATTGATCATCTGATGAGGTTCAAGCTCTGTC 4706  
QY 721 ACTAGCCGCTCACACCAATCTACCTCTGCTTAAATCTCTGGGGGAGATGGTATGCC 780  
DB ACTAGCCGCTCACACCAATCTACCTCTGCTTAAATCTCTGGGGGAGATGGTATGCC 4766  
QY 781 GCCCAATCGCTCCCCCAAGT 840  
DB GCCCAATCGCTCCCCCAAGT 4826  
QY 841 GCTGTGGAGCAATAGGCTTTGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900  
DB GCTGTGGAGCAATAGGCTTTGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4886  
QY 901 GAGGTGAGGCGGCTGT 960  
DB GAGGTGAGGCGGCTGT

DB 4946 GGGGTGGAGCGCGCCTCTGTGCTTTTAAAGTCAATGAGCGGAGATGCCCTTCCACGAG 5005  
QY 961 GACTGTGTTAACTTACTCCCTGCGCAATCTCTCTCTGTGTGCTGTGCTGTGCGGGTCTGT 1020  
DB GACTGTGTTAACTTACTCCCTGCGCAATCTCTCTCTGTGTGCTGTGCTGTGCGGGTCTGT 5006  
QY 1021 TTCGAGCGATATCTGCTGTGCGCATGTGGTCCAGGGAGGGGGCTGTGATGTGATGAAC 1080  
DB TTCGAGCGATATCTGCTGTGCGCATGTGGTCCAGGGAGGGGGCTGTGATGTGATGAAC 5066  
QY 1081 CGGCTGATGCGGTTGGGCTCGGCGGGGTAAACAATTTTCCCGACGACATGATGACAGAC 1140  
DB CGGCTGATGCGGTTGGGCTCGGCGGGGTAAACAATTTTCCCGACGACATGATGACAGAC 5126  
QY 1141 AGCGAGCGCGGAGCATGTGTCACTGATCTCTCGACCTTACATCAACCAACTGTGTG 1200  
DB AGCGAGCGCGGAGCATGTGTCACTGATCTCTCGACCTTACATCAACCAACTGTGTG 5186  
QY 1201 AAGAGGCTTCACAGTGTGATTAACAGAGACTGTCTCAAGGCTGTGCGGCTGTGTGCTTA 1260  
DB AAGAGGCTTCACAGTGTGATTAACAGAGACTGTCTCAAGGCTGTGCGGCTGTGTGCTTA 5246  
QY 1261 AAGAGGCTTCACAGTGTGATTAACAGAGACTGTCTCAAGGCTGTGCGGCTGTGTGCTTA 1320  
DB AAGAGGCTTCACAGTGTGATTAACAGAGACTGTCTCAAGGCTGTGCGGCTGTGTGCTTA 5306  
QY 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGTCCAAAGTGTGTGCAAGGGG 1380  
DB AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGTCCAAAGTGTGTGCAAGGGG 5366  
QY 1381 GTGTGGCGGGAGACGGCATCATGAGACCACTGTCTATGTGTGAGACAGATCACCGGA 1440  
DB GTGTGGCGGGAGACGGCATCATGAGACCACTGTCTATGTGTGAGACAGATCACCGGA 5426  
QY 1441 CATGTCAAAAAGGTTCCATGAGGAATGCTTGGGCTTAAGACCTGTATGATTAACATGTGGCAT 1500  
DB CATGTCAAAAAGGTTCCATGAGGAATGCTTGGGCTTAAGACCTGTATGATTAACATGTGGCAT 5486  
QY 1501 GGAACATTCCTCCATCAACGATATCAACAGGAGCCCTGTGACGCGCTCCACGCGCCAAAC 1560  
DB GGAACATTCCTCCATCAACGATATCAACAGGAGCCCTGTGACGCGCTCCACGCGCCAAAC 5546  
QY 1561 TATTCAAGGGCGCTGT 1620  
DB TATTCAAGGGCGCTGT 5606  
QY 1621 GATTTCCATTAAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
DB GATTTCCATTAAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 5666  
QY 1681 GCGCCCGAATTCCTTCAAGAGGAGTACATTCAGATGCGGCTCAACCAATACCTGTTGGG 1740  
DB GCGCCCGAATTCCTTCAAGAGGAGTACATTCAGATGCGGCTCAACCAATACCTGTTGGG 5726  
QY 1741 AAACTCTCTTACGAGGAGAGTACATTCAGATGCGGCTCAACCAATACCTGTTGGG 1800  
DB AAACTCTCTTACGAGGAGAGTACATTCAGATGCGGCTCAACCAATACCTGTTGGG 5786  
QY 1801 TCGAGACTCCATGCAAGGCCCAACCGGATGATGATGATGATGATGATGATGATGATG 1860  
DB TCGAGACTCCATGCAAGGCCCAACCGGATGATGATGATGATGATGATGATGATGATG 5846  
QY 1861 CCTTCCCATCAACAGAGCGGTAAGCGGCTGAGGCGGAGGAGGCTCTCCCTCTCC 1920  
DB CCTTCCCATCAACAGAGCGGTAAGCGGCTGAGGCGGAGGAGGCTCTCCCTCTCC 5906  
QY 1921 TTGGCAGCTCTTACGCTAGCACTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980  
DB TTGGCAGCTCTTACGCTAGCACTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5966  
QY 1981 CAAAATGACTTCCCAACGCTGATCATGAGGCGCAACTCTCTGTGGGCGGCAATGAGATG 2040  
DB CAAAATGACTTCCCAACGCTGATCATGAGGCGCAACTCTCTGTGGGCGGCAATGAGATG 6026



2041 GGGGGGACATTACCCGCTGAGTCAAGAAACAAGTAGTAATCTTGCATCTTTGCAC 2100  
6086 GGGGGGAATCAACCCGCTGAGTCAAGAAATAAGTAGTAATTTTGCATCTTTGCAG 6145  
2101 CCGCTCGAGCGAGAGAGATGAGCGGAAGTGTCCGTCGCGGAGATCTTGCAGAA 2160  
6146 CCGCTCAAGCGAGAGAGATGAGGAAGTATCGTTCCGCGGAGATCTTGCAGAA 6205  
2161 TCCAGAAATTTCCCAACCAAGATGCCCCGATGCGGCAACGCCCGGATTAACAACCTCCGCTG 2220  
6206 TCCAGAAATTTCCCTGAGCGATGCCCCATATGCGCAACGCCCGGATTAACAACCTCCAG 6265  
2221 CTGAGATCTGGAAGCGCCGAGTACGTCCCTCCAGTGTATCATGGGTGCCCCAGTCCA 2280  
6266 TTRGATCTTGGAAGAGACCCGAGTACGTCTCCAGTGTATCATGGGTGCCCCAGTCCA 6325  
2281 CCTACTAAGACCCCTCTATACCACTCCACGAGAGAGAGACAGTTGTTCTGACAGAA 2340  
6326 CCGCCAAAGCCCTCCGATACCACTCCACGAGAGAGAGAGAGAGTTGTTCTGACAGAA 6385  
2341 TCCACCGTGTCTTCTGCTGCGGAGTTCGCAAAAGCTTTGGTAGCTCCGAGACG 2400  
6386 TCTACCGTGTCTTCTGCTGCGGAGTTCGCAAAAGCTTTGGTAGCTCCGAGACG 6445  
2401 TCGGCGGTGAGACGAGGAGACGAGCAACGCCCTCCAGCAACCTCCGAGAGAGAG 2460  
6446 TCGGCGGTGAGACGAGGAGACGAGCAACGCCCTCCAGCAACCTCCGAGAGAGAG 6505  
2461 GAGAGATCTGAGCGTTGAGTGTATTCCTCCATGCCCCCTTGAAGGAGAGAGAGAGAG 2520  
6506 GGGGAGTCCGAGTGTGATGTATCTCTCCATGCCCCCTTGAAGGAGAGAGAGAGAG 6565  
2521 CCCGATCTCAAGCGAGCGGTCTTGTCTACCGTGAAGAGAGAGAGAGAGAGAGAG 2580  
6566 CCCGATCTCAAGCGAGCGGTCTTGTCTACCGTGAAGAGAGAGAGAGAGAGAGAG 6625  
2581 TCGTGTGATGTCTTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
6626 TCGTGTGATGTCTTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6685  
2641 AGCAAGCTGCCCATCAACGCGTTGAGCAACTCTTGTGCGTCAACCAACATGCTTCA 2700  
6686 ACCAAGCTGCCCATCAACGCGTTGAGCAACTCTTGTGCGTCAACCAACATGCTTCA 6745  
2701 GCTACCAATCCCGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
6746 GCTACCAATCCCGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6805  
2761 ATCCGTGAGATCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
6806 GTCTGTGAG 6865  
2821 AAGGCTAAGCTTCTATCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
6866 AAGGCTAAGCTTCTATCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6925  
2881 TCTAAATTTGGCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
6926 TCTAAATTTGGCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6985  
2941 ATCCGCTCGTGTGAG 3000  
6986 ATCCGCTCGTGTGAG 7045  
3001 ATGCAAAATGAG 3060  
7046 ATGCAAAATGAG 7105  
3061 CTATATGTTTCCAG 3120  
7106 CTATATGTTTCCAG 7165

3121 GTCTCACCCCTCCCTCAGGCTGTGATGAGGCTCTCTGTACAGATTCAGTATCTCTGCA 3180  
7166 GTCTCACCCCTCCCTCAGGCTGTGATGAGGCTCTCTGTACAGATTCAGTATCTCTGCA 7225  
3181 CAGCGAGTGAATCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
7226 CAGCGAGTGAATCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7285  
3241 TATGACACCCGCTGTTTGTGATCAACAGTCACTGAGATGATGATGATGATGATGATGAT 3300  
7286 TATGACACCCGCTGTTTGTGATCAACAGTCACTGAGATGATGATGATGATGATGATGAT 7345  
3301 ATTATCAATGTTGTGATCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
7346 ATTATCAATGTTGTGATCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7405  
3361 CCGCTTATATGAG 3420  
7406 CCGCTTATATGAG 7465  
3421 TGCCTGAG 3480  
7466 TGCCTGAG 7525  
3481 GCTCTGAG 3540  
7526 GCTCTGAG 7585  
3541 GGCCTTGTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
7586 GACCTTGTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7645  
3601 TTCAGGAG 3660  
7646 TTCAGGAG 7705  
3661 GACCTGAG 3720  
7706 GACCTGAG 3765  
3721 AAAAGGTATTAATCTCAACCCGTGAGCC 3749  
7766 AAAAGGTATTAATCTCAACCCGTGAGCC 7794

RESULT 3  
US-09-539-601-10  
; Sequence 10, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; EARLIER FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 7989  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct 1377/NS3-3'/wt  
; NAME/KEY: CDS  
; LOCATION: (342)..(1181)  
; OTHER INFORMATION: hepatitis C virus core-neomycin phosphotransferase  
; OTHER INFORMATION: fusion protein  
; FEATURE:

/ NAME/KEY: RBS  
/ LOCATION: (1190)..(1800)  
/ OTHER INFORMATION: internal ribosome entry site from  
/ OTHER INFORMATION: encephalomyocarditis virus  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (1801)..(7758)  
/ OTHER INFORMATION: hepatitis C virus NS3 - 5B  
/ FEATURE:  
/ NAME/KEY: 3'UTR  
/ LOCATION: (7759)..(7989)  
/ PUBLICATION INFORMATION:  
/ AUTHORS: Lohmann, Volker  
/ AUTHORS: Krner, Frank  
/ AUTHORS: Koch, Jan-Oliver  
/ AUTHORS: Heitman, Ulrike  
/ AUTHORS: Theilmann, Lorenz  
/ AUTHORS: Barrenschiager, Ralf  
/ TITLE: Replication of subgenomic hepatitis C virus RNAs in a  
/ TITLE: hepatoma cell line  
/ JOURNAL: Science  
/ VOLUME: 285  
/ PAGES: 110-113  
/ DATE: 1999-07-02  
/ US-09-539-601-10

Query Match 87.8%; Score 3291.4; DB 4; Length 7989;  
Best Local Similarity .92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 1 TGGAGAGGCGCTTTCACAGGCTCACCCAGTGGATGCCACTTCCGTCCAAACAAAG 60  
DB 3397 TGGAGAGGCGCTTTCACAGGCTCACCCAGTGGATGCCACTTCCGTCCAAACAAAG 3456  
QY 61 CAGGCGAGAGACAACTTCCCTACCTGTCGTCGTAACAGGCTACTGTGTCGCTAGGACC 120  
DB 3457 CAGGCGAGAGACAACTTCCCTACCTGTCGTCGTAACAGGCTACTGTGTCGCTAGGACC 3516  
QY 121 CAGGCCACCTCATCATGAGATCAAAATGTGAAAGTCTCATACGCGTAAAGCTACT 180  
DB 3517 CAGGCTTCACCTCATCATGAGATCAAAATGTGAAAGTCTCATACGCGTAAAGCTACT 3576  
QY 181 CTGCGGGGGGCAACCCCTTGTGTATAGGCTGGAGCCGTCGCAAAACGAGGTACCCCTC 240  
DB 3577 CTGCGGGGGGCAACCCCTTGTGTATAGGCTGGAGCCGTCGCAAAACGAGGTACCCCTC 3636  
QY 241 ACACACCCCTTAACCAATTCATCATGAGATCAAGTGTACACCGACCTGGAGGTCTACG 300  
DB 3637 ACACACCCCTTAACCAATTCATCATGAGATCAAGTGTACACCGACCTGGAGGTCTACG 3696  
QY 301 AGCAGCTGGTGTGTCGGGGGGGCTCTTGACAGCTGTGCTGCGATATGCTTGACACAA 360  
DB 3697 AGCAGCTGGTGTGTCGGGGGGGCTCTTGACAGCTGTGCTGCGATATGCTTGACACAA 3756  
QY 361 GGCAGCGTGTATTTGGTGTAGGATCATTTGTCCGGGCGCCGCGCTATTTGTTCCCGAC 420  
DB 3757 GGCAGCGTGTATTTGGTGTAGGATCATTTGTCCGGGCGCCGCGCTATTTGTTCCCGAC 3816  
QY 421 AGGGAATCTCTTACAGAGAGTTCATGAGATGGAAGAGTGGCGTGCATCTCCCTTAC 480  
DB 3817 AGGGAATCTCTTACAGAGAGTTCATGAGATGGAAGAGTGGCGTGCATCTCCCTTAC 3876  
QY 481 ATGAGAGGGAATGAGCTCGCGAGCAGTCAAGCAAAAGCGCTGGAGTTGCTGTCAG 540  
DB 3877 ATGAGAGGGAATGAGCTCGCGAGCAGTCAAGCAAAAGCGCTGGAGTTGCTGTCAG 3936  
QY 541 ACAGCCACCAAGCAAGCGAGCGCTCTCCGTCGTGAGTCCAAATGGCGAGCCCTT 600  
DB 3937 ACAGCCACCAAGCAAGCGAGCGCTCTCCGTCGTGAGTCCAAATGGCGAGCCCTT 3996  
QY 601 GAGACTCTTGGGGGAAACATGTGAACTTCAACCGGGATACGTAATTAGCAGGC 660  
DB 3997 GAGACTCTTGGGGGAAACATGTGAACTTCAACCGGGATACGTAATTAGCAGGC 4056

QY 661 TTGTCACTCTGCTCGGGAATCCGCGATTGATCACTGATGGCGTTCACAGCTCTGTG 720  
DB 4057 TTGTCACTCTGCTCGGGAATCCGCGATTGATCACTGATGGCGTTCACAGCTCTGTG 4116  
QY 721 ACTAGCCGCTCAACCAATCTACCTCTGCTTAAACATCTGCGGGGATGGGTAGCC 780  
DB 4117 ACTAGCCGCTCAACCAATCTACCTCTGCTTAAACATCTGCGGGGATGGGTAGCC 4176  
QY 781 GCCCACTGCTCCCCAGTCTGCTTCACTTGTAGAGCGCGGCAATTCCTGTGTCG 840  
DB 4177 GCCCACTGCTCCCCAGTCTGCTTCACTTGTAGAGCGCGGCAATTCCTGTGTCG 4236  
QY 841 GCTGTGGAGATAGGCTTGGAAAGTGTGTGACATTTGGCGGGCTATGAGAG 900  
DB 4237 GCTGTGGAGATAGGCTTGGAAAGTGTGTGACATTTGGCGGGCTATGAGAG 4296  
QY 901 GGAGTGGAGGCGGCTCTGTCGCTTAAAGTCAAGCGGGGAATGCCCTCCACCGAG 960  
DB 4297 GGAGTGGAGGCGGCTCTGTCGCTTAAAGTCAAGCGGGGAATGCCCTCCACCGAG 4356  
QY 961 GACTGTGTTAACTTACTCCCTGCAATCTCTCTGTCGTCGTCGTCGTCGTCGTCG 1020  
DB 4357 GACTGTGTTAACTTACTCCCTGCAATCTCTCTGTCGTCGTCGTCGTCGTCGTCG 4416  
QY 1021 TGGCAGCGATACGTGTCGAGCGTGGTCCAGGGAGGGGCTGTGCAATGATGAA 1080  
DB 4417 TGGCAGCGATACGTGTCGAGCGTGGTCCAGGGAGGGGCTGTGCAATGATGAA 4476  
QY 1081 CGGCTGATAGCTTGGCTCGGGGGTAAACATGTTTCCCAACCACTATGTCAGAG 1140  
DB 4477 CGGCTGATAGCTTGGCTCGGGGGTAAACATGTTTCCCAACCACTATGTCAGAG 4536  
QY 1141 AGCAGCGCGCAGCAGTGTCACTCAAGATCTCTCGACTTACTATCAACCACTGTC 1200  
DB 4537 AGCAGCGCGCAGCAGTGTCACTCAAGATCTCTCGACTTACTATCAACCACTGTC 4596  
QY 1201 AAGAGCTTCACCAAGTATTAACAGAGCTCTCCACGCTGCTCGGCTGTGCTGTA 1260  
DB 4597 AAGAGCTTCACCAAGTATTAACAGAGCTCTCCACGCTGCTCGGCTGTGCTGTA 4656  
QY 1261 AAGAGTGTGGGATGAGATATGCAAGTGTGGCTGATTTGAACCTGGCTCCAGTCC 1320  
DB 4657 AAGAGTGTGGGATGAGATATGCAAGTGTGGCTGATTTGAACCTGGCTCCAGTCC 4716  
QY 1321 AAGCTCTGCGGATTTACCGGAGTCCCTTTTCTCATGCAACGTCGGTAAAGGGG 1380  
DB 4717 AAGCTCTGCGGATTTACCGGAGTCCCTTTTCTCATGCAACGTCGGTAAAGGGG 4776  
QY 1381 GTCTGCGGGAGACGGCATATGACAGACCTGCTCATGTGAGACAGATCACCGGA 1440  
DB 4777 GTCTGCGGGAGACGGCATATGACAGACCTGCTCATGTGAGACAGATCACCGGA 4836  
QY 1441 CATGTAAAAAAGTTCATGAGATCTGTGGGCTTAAACCTGTATGTAACATGTGGCAT 1500  
DB 4837 CATGTAAAAAAGTTCATGAGATCTGTGGGCTTAAACCTGTATGTAACATGTGGCAT 4896  
QY 1501 GGAACATTCCTCAACGATACACCAAGGAGCCCTGACAGCCCTCCCAAGGCGCAAC 1560  
DB 4897 GGAACATTCCTCAACGATACACCAAGGAGCCCTGACAGCCCTCCCAAGGCGCAAC 4956  
QY 1561 TATTCAGAGCGCTGTGAGGCTGTGCTGTGAGAGTACGTGAGGTTACGCGGCTGGG 1620  
DB 4957 TATTCAGAGCGCTGTGAGGCTGTGCTGTGAGAGTACGTGAGGTTACGCGGCTGGG 5016  
QY 1621 GATTTCACTAGCTGACAGCATGACCACTGACACATGTAATATCCCGTCCAGGTTCCA 1680  
DB 5017 GATTTCACTAGCTGACAGCATGACCACTGACACATGTAATATCCCGTCCAGGTTCCA 5076  
QY 1681 GCCCGGAATTCCTTCAAGAGTGAATGGGGTGCAGCTGACAGTACGCTCGGCGTGC 1740  
DB 5077 GCCCGGAATTCCTTCAAGAGTGAATGGGGTGCAGCTGACAGTACGCTCGGCGTGC 5136

1741 AAACCTCTCTCAAGGAGAGGTCATTCAGAGTGGGCTCAACCAATACCTGTTGGG 1800  
5137 AAACCCCTCTAGGGAGAGGTCATTCCTGTTGGGCTCAATACCTGTTGGG 5196  
1801 TGGAGCTCCATGCGAGCCCGAAGGATGAGAGTGTCTCACTTCATGCTCAACGAC 1860  
5197 TCACAGCTCCATGCGAGCCCGAAGGATGAGAGTGTCTCACTTCATGCTCAACGAC 5256  
1861 CCTCCCATCATCAGCAGAGAGGCTAAGCGAGGCTGGCGAGGGGGTCTCCCGCTCC 1920  
5257 CCTCCCATCATCAGCAGAGAGGCTAAGCGAGGCTGGCGAGGGGGTCTCCCGCTCC 5316  
1921 TTGGCAGCTCTTCAAGTACGAGTGTCTGCGCTTCTCGAAGGCGCATATCAATACC 1980  
5317 TTGGCAGCTCTTCAAGTACGAGTGTCTGCGCTTCTCGAAGGCGCATATCAATACC 5376  
1981 CAAATGACTTCCAGAGCCTGACCTCATGAGGCGCACTCTGTGGCGGATGAGATG 2040  
5377 CGTCAATGACTTCCAGAGCCTGACCTCATGAGGCGCACTCTGTGGCGGATGAGATG 5436  
2041 GCGCGGAGCATTAACCGCGTGAAGTGAAGAACAGGTATATCTGAGCTCTTCCGAC 2100  
5437 GCGCGGAGCATTAACCGCGTGAAGTGAAGAACAGGTATATCTGAGCTCTTCCGAC 5496  
2101 CCGCTCCAGAGGAGAGATGAGCGGAAGTGTCCGTCGCGGCGAGATCTCGCGGAAA 2160  
5497 CCGCTCCAGAGGAGAGATGAGCGGAAGTGTCCGTCGCGGCGAGATCTCGCGGAAA 5556  
2161 TCCAAAGAAATTCACACAGCAGATGCGGATGCGGATGCGGATTAACAACCTCGCGT 2220  
5557 TCCAGAGAAATTCACACAGCAGATGCGGATGCGGATGCGGATTAACAACCTCGCGT 5616  
2221 CTGGAGTCTGAGAGGCGCGGATGAGTGTCTCCGAGTGTATAGTGGTCCCACTGCGA 2280  
5617 TTAGAGTCTGAGAGGCGCGGATGAGTGTCTCCGAGTGTATAGTGGTCCCACTGCGA 5676  
2281 CCTACTAAGACCCCTCTATACCACTCAAGAGAGAGAGAGAGATGTTCTGACAGAA 2340  
5677 CCTGCAAGGCGCGCTCCGATACCACTCAAGAGAGAGAGAGAGATGTTCTGACAGAA 5736  
2341 TCCACCGTGTCTTCTGCGGAGGCTTCCCAAGAGCTTCTGAGTCTCGAAGCG 2400  
5737 TCTACCGTGTCTTCTGCGGAGGCTTCCCAAGAGCTTCTGAGTCTCGAAGCG 5796  
2401 TGGGCGTGAAGAGCGGAGCGGAGCGGCGCTCTCTGACCAACCTCTCGACAGCGCGA 2460  
5797 TGGGCGTGAAGAGCGGAGCGGAGCGGCGCTCTCTGACCAACCTCTCGACAGCGCGA 5856  
2461 GCGAGATCTGAGTGTGATGATGCTCTCATGCGCGCTTGAAGGGGAGCGGAGGAGC 2520  
5857 GCGGAGTCTGAGTGTGATGATGCTCTCATGCGCGCTTGAAGGGGAGCGGAGGAGC 5916  
2521 CCGCATCTCAGGAGCGGAGCTTGTGTCTACCGTGAAGAGAGCGGAGAGAGCGTCTGC 2580  
5917 CCGCATCTCAGGAGCGGAGCTTGTGTCTACCGTGAAGAGAGCGGAGAGAGCGTCTGC 5976  
2581 TGGCTGCTGATGCTCTCAACATGAGAGAGCGGCTCTGATACCGCATGCGTGGAGGAA 2640  
5977 TGGCTGCTGATGCTCTCAACATGAGAGAGCGGCTCTGATACCGCATGCGTGGAGGAA 6036  
2641 AGCAAGCTCCCATCAAGCGGTTGAGCACTCTTGTGTCGTCACCAACATGCTGAC 2700  
6037 AGCAAGCTCCCATCAAGCGGTTGAGCACTCTTGTGTCGTCACCAACATGCTGAC 2760  
2701 GGTACCATCTCCGAGCGCAAGCGGAGCGGAGAGAGAGTCACTTTGACAGACTGCAA 2760  
6097 GGTACCATCTCCGAGCGCAAGCGGAGCGGAGAGAGAGTCACTTTGACAGACTGCAA 6156  
2761 ATCTGAGAGATCACTACAGAGAGCGTCTCAAGAGAGATGAGCGGAGCGTCTCAAGT 2820  
6157 GTCTGAGAGATCACTACAGAGAGCGTCTCAAGAGAGATGAGCGGAGCGTCTCAAGT 6216  
2821 AAGGCTAAGCTTCTATCACTAGAGAGAGCGTCAAGCTGACGCGCCCATTCGCGCAA 2880

6217 AAGGCTAAGCTTCTATCCGTGAGAGAGCGCTTAAGCTGACCGCCCATTCGCGCAA 6276  
2881 TCTAAATTTGGCTATGAGGCAAGAGCGTCCGAAACCTTCCAGAGAGCGCTTAACAC 2940  
6277 TCTAAATTTGGCTATGAGGCAAGAGCGTCCGAAACCTTCCAGAGAGCGCTTAACAC 6336  
2941 ATCCGCTCCGTGAGAGAGCTTGTGAGAGACCTGAGAAACCAATTTGACACACATC 3000  
6337 ATCCGCTCCGTGAGAGAGCTTGTGAGAGACCTGAGAAACCAATTTGACACACATC 6396  
3001 ATGGCAAAATGAGGTTTCTGCGTCAACAGAGAGAGAGCGGAGCGGAGCGTCCG 3060  
6397 ATGGCAAAATGAGGTTTCTGCGTCAACAGAGAGAGAGCGGAGCGGAGCGTCCG 6456  
3061 CTATGCTGTTCCAGACTTGGGAGGCTCGTGTGTGAGAGAAATGAGGCTCTATGACG 3120  
6457 CTATGCTGTTCCAGACTTGGGAGGCTCGTGTGTGAGAGAAATGAGGCTCTATGACG 6516  
3121 GTCTCAACCTCCCTCAGGCTGTGATGAGGCTCTGTGACGATTCAGATTCCTCTGGA 3180  
6517 GTCTCAACCTCCCTCAGGCTGTGATGAGGCTCTGTGACGATTCAGATTCCTCTGGA 6576  
3181 CAGCGGCTGAGTCTCTGAGTGAACGCTGGAATCAAGAGAGAGAGCGGAGCGTTCGA 3240  
6577 CAGCGGCTGAGTCTCTGAGTGAACGCTGGAATCAAGAGAGAGAGCGGAGCGTTCGA 6636  
3241 TATGACACCGCTGTTTGAATCAACAGTCACTGAGATGATCCGTGTAGAGAGTCA 3300  
6637 TATGACACCGCTGTTTGAATCAACAGTCACTGAGATGATCCGTGTAGAGAGTCA 6696  
3301 ATTTATCAATGTTGATGATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3360  
6697 ATTTATCAATGTTGATGATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6756  
3361 CCGCTTTATATGAGGAGGCTCCCTGATTAATTAAGAGGAGGAGGAGGAGGAGGAGG 3420  
6757 CCGCTTTATATGAGGAGGCTCCCTGATTAATTAAGAGGAGGAGGAGGAGGAGGAGG 6816  
3421 TGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480  
6817 TGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6876  
3481 GCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3540  
6877 GCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6936  
3541 GGCCTTGTCTATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3600  
6937 GGCCTTGTCTATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6996  
3601 TTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3660  
6997 TTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7056  
3661 GACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3720  
7057 GACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7116  
3721 AAAAGGTTACTACCTCAACCGGAGCC 3749  
7117 AAAAGGTTACTATCTCAACCGGAGCC 7145

RESULT 4  
US-09-539-601-7  
Sequence 7, Application US/09539601C  
Patent No. 6630343  
GENERAL INFORMATION:  
APPLICANT: Barteneschlager, Ralf FW  
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
FILE REFERENCE: all sequences  
CURRENT APPLICATION NUMBER: US/09/539,601C

CURRENT FILING DATE: 2001-08-30
EARLIER APPLICATION NUMBER: 139 15 178.4 GERMANY
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 8001
TYPE: DNA
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)..(341)
OTHER INFORMATION: construct I389/NS3-3'/wt
NAME/KEY: CDS
LOCATION: (342)..(1193)
OTHER INFORMATION: hepatitis C virus core-neomycin
OTHER INFORMATION: phosphotransferase fusion protein
FEATURE:
NAME/KEY: RBS
LOCATION: (1202)..(1812)
OTHER INFORMATION: internal ribosome entry site from
OTHER INFORMATION: encephalomyocarditis virus
FEATURE:
NAME/KEY: CDS
LOCATION: (1813)..(7770)
OTHER INFORMATION: hepatitis C virus nonstructural proteins NS3-5B
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (7771)..(8001)
PUBLICATION INFORMATION:
AUTHORS: Lohmann, Volker
AUTHORS: Krner, Frank
AUTHORS: Koch, Jan-Oliver
AUTHORS: Herian, Ulrike
AUTHORS: Theilmann, Lorenz
AUTHORS: Bartenschlager, Ralf
TITLE: Replication of subgenomic hepatitis c virus RNAs in a
TITLE: hepatoma cell line
JOURNAL: Science
VOLUME: 285
PAGES: 110-113
DATE: 1999-07-02
US-09-539-601-7

Query Match 87.8%; Score 3291.4; DB 4; Length 8001;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

1 TGAGAGGCGCTTTCACAGGCTCACCAGTGATGCCACTTCTGTCCTCCAAACAAAG 60
3409 TGAGAGGCGCTTTCACAGGCTCACCAGTGATGCCACTTCTGTCCTCCAAACAAAG 3468
61 CAGGACAGAGCAACATTTCCCTTACCTGTGGGCTACAGGCTACTGTGCGCTAGGGCC 120
3469 CAGGACAGAGCAACATTTCCCTTACCTGTGGGCTACAGGCTACTGTGCGCTAGGGCC 3528
121 CAGGCGCCACTCTCATATGAGATCAATGTGAAAGTGTCTCATACGGCTTAAAGCTTACT 180
3529 CAGGCTTCACCTCATGTGGAGCAAAATGTGAAAGTGTCTCATACGGCTTAAAGCTTACT 3588
181 CTGGCGGGGCAACACCTTGTGTATAGGCTGGAGCGCTCCAAACGAGTCAACCTC 240
3589 CTGGCGGGGCAACACCTTGTGTATAGGCTGGAGCGCTCCAAACGAGTCAACCTC 3648
241 ACAACCCCATTAACCAATTCATCATGATGATGATGATGATGATGATGATGATGATGAT 300
3649 ACAACCCCATTAACCAATTCATCATGATGATGATGATGATGATGATGATGATGATGAT 3708
301 AGACCTGGGT 360
3709 AGACCTGGGT 3768

QY 361 GGCAGCGGTGATCTTGTGGTAGATCATCTTGTCCGGGCGGCGCTATTGTCCTCCGAC 420
DB 3769 GGCAGCGGTGATCTTGTGGTAGATCATCTTGTCCGGGCGGCGCTATTGTCCTCCGAC 3828
QY 421 AGGAGAGCTTCTTACAGGAGTTCATGATGAGAGAGTGGCGGTGACCTCCCTTAC 480
DB 3829 AGGAGAGCTTCTTACAGGAGTTCATGATGAGAGAGTGGCGGTGACCTCCCTTAC 3888
QY 481 ATCGACAGGAGATCAGCTCCGAGCAGTTCAGCAAAACGCTCGGTTGTGCGAG 540
DB 3889 ATCGACAGGAGATCAGCTCCGAGCAGTTCAGCAAAACGCTCGGTTGTGCGAG 3948
QY 541 AAGGACCAAGCAAGCGAGGCGCTGCTCCGTTGTGAGTCCAAAGTGGAGCCCTT 600
DB 3949 AAGGACCAAGCAAGCGAGGCGCTGCTCCGTTGTGAGTCCAAAGTGGAGCCCTT 4008
QY 601 GAGACCTTGTGGCGCAACATGTGGAATCTTCAAGCGGAGTACAGTACTTACAGGC 660
DB 4009 GAGACCTTGTGGCGCAACATGTGGAATCTTCAAGCGGAGTACAGTACTTACAGGC 4068
QY 661 TTGTCCACTCTGCTGGGAATCCCGATTTGATCATGATGATGATGATGATGATGATGAT 720
DB 4069 TTGTCCACTCTGCTGGGAATCCCGATTTGATCATGATGATGATGATGATGATGATGAT 4128
QY 721 ACTAGCGGCTCACCAACCAATCTACCTCTGTTAACATCTCTGGGGGATGGGTAGCC 780
DB 4129 ACTAGCGGCTCACCAACCAATCTACCTCTGTTAACATCTCTGGGGGATGGGTAGCC 4188
QY 781 GCCCAACTGCTCCCCGAGTCTGCTTCAAGTTCGTAGGGGCGGCAATTCCTGCTGTCG 840
DB 4189 GCCCAACTGCTCCCCGAGTCTGCTTCAAGTTCGTAGGGGCGGCAATTCCTGCTGTCG 4248
QY 841 GCTGTGGAGCATAGGCTTGGGAAAGTGTGTGACATCTTGGCGGGCTATGAGAGA 900
DB 4249 GCTGTGGAGCATAGGCTTGGGAAAGTGTGTGACATCTTGGCGGGCTATGAGAGA 4308
QY 901 GGAATGGAGCGGCGCTGCTGAGCTTTAAAGTTCATGAGCGGGAATGCCCTTCACCGAG 960
DB 4309 GGAATGGAGCGGCGCTGCTGAGCTTTAAAGTTCATGAGCGGGAATGCCCTTCACCGAG 4368
QY 961 GACCTGGTTAACTTCTCCCTGCAATCCCTCTCTGCTGGTGGCTGGGGGTCGAG 1020
DB 4369 GACCTGGTTAACTTCTCCCTGCAATCCCTCTCTGCTGGTGGCTGGGGGTCGAG 4428
QY 1021 TGCGAGCATACTGCTGGGACGCGAGTGGTCCAGGGGAGGGGCTGTGAGTGAATGAAAC 1080
DB 4429 TGCGAGCATACTGCTGGGACGCGAGTGGTCCAGGGGAGGGGCTGTGAGTGAATGAAAC 4488
QY 1081 CGGCTGATAGCGTTGCTGCGGGGGTAAACAATGTTTCCCAACGCACTATGTGCCAG 1140
DB 4489 CGGCTGATAGCGTTGCTGCGGGGGTAAACAATGTTTCCCAACGCACTATGTGCCAG 4548
QY 1141 AGGAGCGCGGACAGTGTCACTGATCTCTCCGACTTATCTACCACTATGACCACTGTTG 1200
DB 4549 AGGAGCGCGGACAGTGTCACTGATCTCTCTCCGACTTATCTACCACTATGACCACTGTTG 4608
QY 1201 AAGAGGCTCCACAGTGAATTAACGAGACTGCTCCACGCGCTGCTCGGCTGTGGCTTA 1260
DB 4609 AAGAGGCTCCACAGTGAATTAACGAGACTGCTCCACGCGCTGCTCGGCTGTGGCTTA 4668
QY 1261 AAGGATGTTTGGAGATGATATGACAGTGTGCTGCTTCAAGACTGCTGCTCCAGTCC 1320
DB 4669 AAGGATGTTTGGAGATGATATGACAGTGTGCTGCTTCAAGACTGCTGCTCCAGTCC 4728
QY 1321 AAGCTCTGCGGAGATTAACGGAGTCCCTTTTCTCATGCAACGTTGGTGAACAAGGGG 1380
DB 4729 AAGCTCTGCGGAGATTAACGGAGTCCCTTTTCTCATGCAACGTTGGTGAACAAGGGG 4788
QY 1381 GTCTGGGGGAGAGCGGATCATGACAGCACTGCTCATGTGAGACACAGATCACCGGA 1440
DB 4789 GTCTGGGGGAGAGCGGATCATGACAGCACTGCTCATGTGAGACACAGATCACCGGA 4848
QY 1441 CATGTCAAAAACGGTTTCATGAGGATGTTGGGCTTAAGACCTGTATGAATGATGGCAT 1500

Db 4849 CATTGAAAAAGCGTTCCATGAGAGATCTGGGGCTTAGAGACTGTGTATGTAACAAGTGGCAT 4908  
Qy 1501 GGAACATTTCCCATCAACCATACCAAGGAGCCCTTCAGAGCCCTCCCAAGGCCAAC 1560  
Db 4909 GGAACATTTCCCATCAACCATACCAAGGAGCCCTTCAGAGCCCTCCCAAGGCCAAC 4968  
Qy 1561 TATTCAGAGGCGGTGGGGGGGTGGCTGAGAGAGTGAAGGTTACGGGGGTGGG 1620  
Db 4969 TATTCAGAGGCGGTGGGGGGGTGGCTGAGAGAGTGAAGGTTACGGGGGTGGG 5028  
Qy 1621 GATTTCCACTACGTGACGAGCATGACCACTGACAGGTAAATGCCGTGACAGTTTCA 1680  
Db 5029 GATTTCCACTACGTGACGAGCATGACCACTGACAGGTAAATGCCGTGACAGTTTCCG 5088  
Qy 1681 GCGCCCGAATTTCTTCAAGAAATGTGATGGGTGGCTGACAGGTACCGTCCGGGTGC 1740  
Db 5089 GCGCCCGAATTTCTTCAAGAAATGTGATGGGTGGCTGACAGGTACCGTCCGGGTGC 5148  
Qy 1741 AAACCTCTCTACGGGAGAGGTACATTCAGGTGGGCTCAACCAATTCGTGTGG 1800  
Db 5149 AAACCTCTCTACGGGAGAGGTACATTCAGGTGGGCTCAATCAATTCGTGTGG 5208  
Qy 1801 TCGAGCTCCATGCGAGCCGGAACCGAGTGTAGAGTGTCTCACTTCATGCTACCGAC 1860  
Db 5209 TCGAGCTCCATGCGAGCCGGAACCGAGTGTAGAGTGTCTCACTTCATGCTACCGAC 5268  
Qy 1861 CCCTCCCATCACTACAGAGAGAGCGGTAAAGCGAGGTGGCCAGGGGGTCTCCCTCC 1920  
Db 5269 CCCTCCCATCACTACAGAGAGAGCGGTAAAGCGAGGTGGCCAGGGGGTCTCCCTCC 5328  
Qy 1921 TTGGGAGGCTTTCACTAGAGAGTGTCTGGGCTTCTCCGAGGCGACATCAATTAC 1980  
Db 5329 TTGGGAGGCTTTCACTAGAGAGTGTCTGGGCTTCTCCGAGGCGACATCAATTAC 5388  
Qy 1981 CAAATGACTTCCAGAGCGTGAACCTCATGAGCGCAACCTCTGTGCGGCGATGAGAT 2040  
Db 5389 CCAATGACTTCCAGAGCGTGAACCTCATGAGCGCAACCTCTGTGCGGCGAGATG 5448  
Qy 2041 GGGGGGAGCATTAACCGCGGTGAGTGAAGAGCAAGATGATTCCTGATCTTTTGAC 2100  
Db 5449 GGGGGGAGCATTAACCGCGGTGAGTGAAGAGCAAGATGATTCCTGATCTTTTGAC 5508  
Qy 2101 CCGCTCCAGCGGAGAGAGTGAAGCGGAAATGTCCTCCGCGGAGATCCGCGGAA 2160  
Db 5509 CCGCTCCAGCGGAGAGAGTGAAGCGGAAATGTCCTCCGCGGAGATCCGCGGAG 5568  
Qy 2161 TCCAGAAATTTCCACAGAGATGCGGATGGGAGCGCCGAGTTTAAACCTTCGCTG 2220  
Db 5569 TCCAGAAATTTCCCTGAGAGATGCGGATGGGAGCGCCGAGTTTAAACCTTCGAC 5528  
Qy 2221 CTGAGTCTGAGAGGCGCGGAGTACGTCCTCCAGTGTACATGAGTCCCATGCA 2280  
Db 5629 TTAGAGTCTTGGAGAGAGCGGAGTACGTCCTCCAGTGTACATGAGTCCCATG 5688  
Qy 2281 CCTACTAAGACCCCTCTATACCACTCAACGAGAGAGAGCAGTGTTCGACAGAA 2340  
Db 5689 CCTACTAAGACCCCTCTATACCACTCAACGAGAGAGAGCAGTGTTCGACAGAA 5748  
Qy 2341 TCCACGCTGTCTTCTGCGGAGAGTGGCCACAAAGCTTTCGATGCTCGAACG 2400  
Db 5749 TCTACGCTGTCTTCTGCGGAGAGTGGCCACAAAGCTTTCGATGCTCGAACG 5808  
Qy 2401 TCGGCGGTGAGAGCGGAGCGGAGCGGCGCTCTCTGACCAACCTTCGAGCGGAG 2460  
Db 5809 TCGGCGGTGAGAGCGGAGCGGAGCGGCGCTCTCTGACCAACCTTCGAGCGGAG 5868  
Qy 2461 GAGAGTCTGAGAGTGTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATG 2520  
Db 5869 GAGAGTCTGAGAGTGTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATG 5928  
Qy 2521 CCGGATCTGAGAGCGGAGTCTTGGTCTACCGTGTGAGAGCGGAGTGTGAGAGTCT 2580  
Db 5929 CCGGATCTGAGAGCGGAGTCTTGGTCTACCGTGTGAGAGCGGAGTGTGAGAGTCT 5988  
Qy 2581 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
Db 5989 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6048  
Qy 2641 AGCAAGTCCCATCAACCGGTTGAGCAACTCTTGTCTGCTGCTGCTGCTGCTGCT 2700  
Db 6049 ACCAAGTCCCATCAACCGGTTGAGCAACTCTTGTCTGCTGCTGCTGCTGCTGCT 6108  
Qy 2701 GCTTACCAATCCCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 2760  
Db 6109 GCTTACCAATCCCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 6168  
Qy 2761 ATCCGAGAGTACATCAACGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 2820  
Db 6169 ATCCGAGAGTACATCAACGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 6228  
Qy 2821 AAGGCTAAGCTTCTATCACTAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 2880  
Db 6229 AAGGCTAAGCTTCTATCACTAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 6288  
Qy 2881 TCTAAATTTGGCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940  
Db 6289 TCTAAATTTGGCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6348  
Qy 2941 ATCCGCTCCGTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3000  
Db 6349 ATCCGCTCCGTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6408  
Qy 3001 ATGGCAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3060  
Db 6409 ATGGCAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6468  
Qy 3061 CTATGATGATCCAGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3120  
Db 6469 CTATGATGATCCAGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6528  
Qy 3121 GTCTCAACCTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3180  
Db 6529 GTCTCAACCTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6588  
Qy 3181 CAGCGGAGTGAAGTCTGAGTGAAGCGGAGTGAAGCGGAGTGAAGCGGAGTGAAG 3240  
Db 6589 CAGCGGAGTGAAGTCTGAGTGAAGCGGAGTGAAGCGGAGTGAAGCGGAGTGAAG 6648  
Qy 3241 TATGACACCGGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTGAGAGTCA 3300  
Db 6649 TATGACACCGGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTGAGAGTCA 6708  
Qy 3301 ATTATCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360  
Db 6709 ATTATCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6768  
Qy 3361 CCGCTTATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3420  
Db 6769 CCGCTTATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6828  
Qy 3421 TGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3480  
Db 6829 TGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6888  
Qy 3481 GCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3540  
Db 6889 GCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6948  
Qy 3541 GAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600  
Db 6949 GAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7008  
Qy 3601 TTCAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660  
Db 7009 TTCAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7068

Oy		3661	GACCTGAGAGTTGATTAACATCATGTCTCTCCAAATGTGTGGCGCAGATGCATCTGGC	3720
Dd		7069	GACTTGAGGTTGATTAACAATCAATGCTCTCCCAATGTGTCAAGTGGCGCAGATGCATCTGGC	7128
Oy		3721	AAAAGGTAATACTCACTCACCCCGTAGACC	3749
Dd		7129	AAAAGGTAATACTCACTCACCCCGTAGACC	7157
		RESULT 5		
		US-09-539-601-4		
		; Sequence 4, Application US/09539601C		
		Patent No. 6630343		
		GENERAL INFORMATION:		
		APPLICANT: Bartenechlagier, Ralf FW		
		TITLE OF INVENTION: Hepatitis C Virus Cell Culture System		
		FILE REFERENCE: all sequences		
		CURRENT APPLICATION NUMBER: US/09/539,601C		
		CURRENT FILING DATE: 2001-08-30		
		EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY		
		EARLIER FILING DATE: 1999-04-03		
		NUMBER OF SEQ ID NOS: 51		
		SOFTWARE: PatentIn Ver. 2.1		
		SEQ ID NO 4		
		LENGTH: 8637		
		TYPE: DNA		
		ORGANISM: Hepatitis C virus		
		FEATURE:		
		NAME/KEY: 5'UTR		
		LOCATION: (1) ..(341)		
		OTHER INFORMATION: construct 1377/NS2-3'/wt		
		FEATURE:		
		NAME/KEY: CDS		
		LOCATION: (342)..(1181)		
		OTHER INFORMATION: HCV core-neomycin phosphotransferase fusion		
		OTHER INFORMATION: protein		
		FEATURE:		
		NAME/KEY: RBS		
		LOCATION: (1190)..(1800)		
		OTHER INFORMATION: internal ribosome entry site from		
		OTHER INFORMATION: encephalomyocarditis virus		
		FEATURE:		
		NAME/KEY: CDS		
		LOCATION: (1801)..(8406)		
		OTHER INFORMATION: hepatitis C virus NS2 - 5B		
		FEATURE:		
		NAME/KEY: 3'UTR		
		LOCATION: (8407)..(8637)		
		PUBLICATION INFORMATION:		
		AUTHORS: Lohmann, Volker		
		AUTHORS: Krner, Frank		
		AUTHORS: Koch, Jan-Oliver		
		AUTHORS: Herian, Ulrike		
		AUTHORS: Theilmann, Lorenz		
		AUTHORS: Bartenschlager, Ralf		
		TITLE: Replication of subgenomic hepatitis c virus RNAs in a		
		TITLE: hepatoma cell line		
		JOURNAL: Science		
		VOLUME: 285		
		PAGES: 110-113		
		DATE: 1999-07-02		
		US-09-539-601-4		
Query Match		87.8%;	Score 3291.4;	DB 4; Length 8637;
Best Local Similarity		92.4%;	Pred. No. 0;	
Matches 3463;		Conservative 0;	Mismatches 286;	Indels 0; Gaps 0;
Oy		1	TGGAGAGGCGCTTCACAGAGGCTCACGCCAGTGATGCCCACTTCTCTGCCCCAACAAGAAG	60
Dd		4045	TGGAGAGGCGCTTTCACAGGCTCACGCCACATAGAGAGCCCAATTCTTGTGCCAGACTAAG	4104
Oy		61	CAGGAGAGACACACTTCCCCTAACCTCGTGGCGTACAGAGCTACTGTGTGCGCTAAGGCC	120

Db	4105	CAGGCAAGAAACAACTTCCCTTACCTGTGACATCCAGGCTACGGTGTGCGCAGGGCT	4164
OY	121	CAGGCCCCACCTCCATCATGGAGATCAAAATGTGAAAGTGTCTCATACGGCTAAAGCTTACT	180
Db	4165	CAGGCTCCACTCCATCATGTGTGGAAACAAATGTGAAAGTGTCTCATACGGCTAAAGCTTACG	4224
OY	181	CTGCGCGGGCCCAACCCCTTGTGTGTATAGCTGTGGAGCCCTCCAAAACAGAGTCAACCTC	240
Db	4225	CTGCACCGGGCCCAACGCCCTGTGTGTATAGCTGTGGAGCCCTCCAAAACAGAGTTTACTAC	4284
OY	241	ACACACCCCAATTAACCAAAATTCATCATATGAGGACATGTCACGCCGACTGTGAGAGTGTACG	300
Db	4285	ACACACCCCAATTAACCAAAATTCATCATATGAGGACATGTCACGCCGACTGTGAGAGTGTACG	4344
OY	301	AGCAGCTTGTGTGTGTGTGGCGGGGTCTTGCAGACTTGTGCTGTGATTTGCTTGACAA	360
Db	4345	AGCAGCTTGTGTGTGTGTGGCGGGGTCTTGCAGACTTGTGCTGTGATTTGCTTGACAA	4404
OY	361	GGCAGCGTGTCTATTGTGTGGTATAGATATCTTGTTCGGGGCGGCGGCTATTTGTTCCCGAC	420
Db	4405	GGCAGCGTGTGTCTATTGTGTGGCGAGATATCTTGTTCGGGAAAGCGGGCATCATTTCCGAC	4464
OY	421	AGGGAAGTCTCTTACCGAGAGTTGTGATGAGATGTGAAAGATGCGGGTGTGCACCTCCCTTAC	480
Db	4465	AGGGAAGTCTCTTACCGGAGTTGTGATGAGATGTGAAAGATGCGGGTGTGCACCTCCCTTAC	4524
OY	481	ATCGAGAGGGAATGACAGCTGCGCCGACAGTTCAAGCAAAAAGGCTTCGGGTGTCTGACG	540
Db	4525	ATCGAGAGGGAATGACAGCTGCGCCGACAGTTCAAGCAAAAAGGCAATTCGGGTGTCTGCAA	4584
OY	541	ACAGCCACCAAGCAGCGAGGCGCGTCTCCGTGTGTGAGTTCAGATGTGCCAGCCCTT	600
Db	4585	ACAGCCACCAAGCAGCGAGGCGCGTCTCCGTGTGTGAGTTCAGATGTGCCAGCCCTC	4644
OY	601	GAGACCTTCTGGGGGGAACAATGTGGAACCTTATCAGCGGGATACAGTACTTGACAGGC	660
Db	4645	GAGACCTTCTGGGGGGAACATATGTGAATTTATCAGCGGGATACATATTTGACAGGC	4704
OY	661	TTGTTCACCTGTGCGTGGGAATCCCGCATTTGATGCATCACTGATGGGTTTACAGCCTGTGC	720
Db	4705	TTGTTCACCTGTGCGTGGGAATCCCGCATTTGATGCATCACTGATGGGTTTACAGCCTTATC	4764
OY	721	ACTNAGCCCGCTACCAACCCAAATCTACCTCTCTGTCTTAAATCTCTGGGGGAGATGGGTAGCC	780
Db	4765	ACCAGCCCGCTCACACCCCAACATACCTCTCTGTCTTAAATCTCTGGGGGAGATGGGTAGCC	4824
OY	781	GCCCAACTCGCTCCCGCAGTGTGCTTCAAGCTTTCGATGGCGCGCGGATGTGCTGTGCG	840
Db	4825	GCCCAACTCGCTCTCCAGCGCTGTCTTCTCTTTCGATGGCGCGCGGATGTGCTGTGAGCG	4884
OY	841	GCTGTGTGACATAGGCTTGTGGGAAGTGTCTTGACATCTTGTGGCGGCTATGAGCA	900
Db	4885	GCTGTGTGACATAGGCTTGTGGGAAGTGTCTTGAGATTTTGTGGCAGGTTATGAGCA	4944
OY	901	GGAAGTGTGAGCGCGCTGTGTGGCTTTTAAAGTCAAGCGCGCAATATGCTTCAACGAG	960
Db	4945	GGAAGTGTGAGCGCGCTGTGTGGCTTTTAAAGTCAAGCGCGCAATATGCTTCAACGAG	5004
OY	961	GACCTGTGTAACTTACTCCCTTGACATCTCTCTGTGTCGTCGAGTGTGCGGGGTGTG	1020
Db	5005	GACCTGTGTAACTTACTCCCTGTGTATCTCTCTGTGTCGTCGAGTGTGCGGGGTGTG	5064
OY	1021	TGCGCAGGAACTAGCGTGTGCGCAGTGTGGTTCAGAGGAGGGGCGCTGTGACATGATGAAC	1080
Db	5065	TGCGCAGGAACTAGCGTGTGCGCAGTGTGGCCCAAGGAGAGGGGCGCTGTGACATGATGAAC	5124
OY	1081	CGGCTGATAGGCTTGTGCTGTGGCGGGGTAAACATGTTTCCCAACGCACTATGTGTACAG	1144
Db	5125	CGGCTGATAGGCTTGTGCTGTGGCGGGGTAAACATGTTTCCCAACGCACTATGTGTACAG	5184
OY	1141	AGCGACCGCGAGACGTGTCACTCAATCTCTTCCAGCTTATCATACCAACTGTGTG	1200

D 5185 AGCAGCGCTGCAGACGCTGCTCACTAGATCTCTAGTCTTAACATCACTAGCTGCTG 5244  
Q 1201 AAGAGGCTCCACCAAGTGAATTAAGAGAGCTGCTCCAGCGCTGCTCGGGCTGTTGAGCTA 1260  
D 5245 AAGAGGCTTACCAAGTGAATTAAGAGAGCTGCTCCAGCGCTGCTCGGGCTGTTGAGCTA 5304  
Q 1261 AAGAGTGTGGAGCTGATATGACAGATTTGGCTGATCTTCAAGACTGGCTTCACTGTC 1320  
D 5305 AGAGATGTTGGATTTGATATGACAGGTTGATGATTTCAAGACTGGCTTCACTGTC 5364  
Q 1321 AAGCTCTGCGGAGTATACCGGAGTCCCTTTTCTCATGCCAAGCTGGGTACAAAGGG 1380  
D 5365 AAGCTCTGCGGAGTATGCGGAGTCCCTTTCTCATGTCAACCTGGGTACAAAGGG 5424  
Q 1381 GTCGCGGGGAGACCGGATCATGACAGACCGCTCATGTGGAGACAGATCAACCGA 1440  
D 5425 GTCTGGCGGGGAGACCGGATCATGACAGACCGCTCATGTGGAGACAGATCAACCGA 5484  
Q 1441 CATGTCAAAAACGGTTCCATGAGATCGTTGGGCTTAAGACTGTAGTAAATATGTGAT 1500  
D 5485 CATGTCAAAAACGGTTCCATGAGATCGTTGGGCTTAAGACTGTAGTAAATATGTGAT 5544  
Q 1501 GGAACATTTCCCATCAACGATACCAACGAGCGCCCTGCAAGCCCTCCAGCGCAAC 1560  
D 5545 GGAACATTTCCCATCAACGATACCAACGAGCGCCCTGCAAGCCCTCCAGCGCAAC 5604  
Q 1561 TATTCAGGCGGCTGTGGCGGGTGTGCTGTGAGAGTACGTGAGATTAAGCGGGTGGG 1620  
D 5605 TATTCAGGCGGCTGTGGCGGGTGTGCTGTGAGAGTACGTGAGATTAAGCGGGTGGG 5664  
Q 1621 GATTTCCACTACGTGACGAGATGACCACTGACAAAGTAAATGCGGCTCCAGTTTCA 1680  
D 5665 GATTTCCACTACGTGACGAGATGACCACTGACAAAGTAAATGCGGCTCCAGTTTCC 5724  
Q 1681 GCCCCGAATTTCTTCAAGAGTGAATGGGCTGCGCTGCAAGTACGCTCCGCGTGC 1740  
D 5725 GCCCCGAATTTCTTCAAGAGTGAATGGGCTGCGCTGCAAGTACGCTCCGCGTGC 5784  
Q 1741 AATCTCTCTTACGAGAGAGTCACTTCCAGTGGGCTCAACCAATACCTGTGGTGG 1800  
D 5785 AATCTCTCTTACGAGAGAGTCACTTCCAGTGGGCTCAACCAATACCTGTGGTGG 5844  
Q 1801 TGCGAGCTCCCATGCGAGCGGCAACCGGATGTAGAGTGTCACTTCCATGCTCACCGAC 1860  
D 5845 TGCGAGCTCCCATGCGAGCGGCAACCGGATGTAGAGTGTCACTTCCATGCTCACCGAC 5904  
Q 1861 CCCTCCCATCAACGAGAGAGCGGCTAAGCGAGGCTGCGCAGGGGCTTCCCTCCCTCC 1920  
D 5905 CCCTCCCATCAACGAGAGAGCGGCTAAGCGAGGCTGCGCAGGGGCTTCCCTCCCTCC 5964  
Q 1921 TTGGCCAGCTCTTACGCTAGCCAGTTGTGTGCGCTTCTCTGAAGCGGACATACATTAAC 1980  
D 5965 TTGGCCAGCTCTTACGCTAGCCAGTTGTGTGCGCTTCTCTGAAGCGGACATACATTAAC 6024  
Q 1981 CAAATGAGCTTCCCAAGCGCTGACCTCATGAGGCGCAACCTCGTGGCGGCAATGAGATG 2040  
D 6025 CGTCAATGAGCTTCCCAAGCGCTGACCTCATGAGGCGCAACCTCGTGGCGGCAATGAGATG 6084  
Q 2041 GCGCGGAGCATTTACCGCGTGAAGTACAGAGAAAGGTAGTAACTCTGATCTTTTCGAC 2100  
D 6085 GCGCGGAGCATTTACCGCGTGAAGTACAGAGAAAGGTAGTAACTCTGATCTTTTCGAC 6144  
Q 2101 CCGCTCCAGCGGAGAGATGAGCGGAAAGTGTCCGTCCGCGCGAGATCTTGGCGAAA 2160  
D 6145 CCGCTCCAGCGGAGAGATGAGCGGAAAGTGTCCGTCCGCGCGAGATCTTGGCGAGG 6204  
Q 2161 TCCAAAGAAATTCACACAGGATGCGCGCATGAGGCAACCGGATTAACAACCTCCGCTG 2220  
D 6205 TCCAAAGAAATTCACACAGGATGCGCGCATGAGGCAACCGGATTAACAACCTCCGCTG 6264  
Q 2221 CTGAGATCTCTGAAGAGCGCGGAGTACGTCTCTCAATGTGTACATGGGTCCCACTGCA 2280  
D 6265 TTAGATCTCTGAAGAGCGCGGAGTACGTCTCTCAATGTGTACATGGGTCCCACTGCGG 6324

Q 2281 CTTACTAAGACCCCTCTTATACCACTCCACGAGAGAGAGACAGTGTGTTGACAGAA 2340  
D 6325 CTTGCAAGGCGCCCTCCGATACCACTCCACGAGAGAGAGAGAGTGTGTTGTTGACAGAA 6384  
Q 2341 TCCACCGTGTCTTTCGCTGCGGAGTGTGCGCAAAAGCTTTCGATGCTCCGAACCG 2400  
D 6385 TCTAACCGTGTCTTTCGCTGCGGAGTGTGCGCAAAAGCTTTCGAGCTCCGAATCG 6444  
Q 2401 TCGGCGGTGACAGCGGAGACGAGCAACGCGCCCTCCGACCAACCTCCGACGAGGAG 2460  
D 6445 TCGGCGGTGACAGCGGAGACGAGCAACGCGCCCTCCGACCAACCTCCGACGAGGAG 6504  
Q 2461 GCAGATCTGACCTTGAATGCTATCTCTCATATGCTCCCTTGAAGGAGGAGCGGAGGAG 2520  
D 6505 GCGGATCTGACCTTGAATGCTATCTCTCATATGCTCCCTTGAAGGAGGAGCGGAGGAG 6564  
Q 2521 CCGGATCTGACCGAGCGGCTTGTGTCTTACCGTGAATGAGAGGCGGAGGAGGAGG 2580  
D 6565 CCGGATCTGACCGAGCGGCTTGTGTCTTACCGTGAATGAGAGGCGGAGGAGGAGGAGG 6624  
Q 2581 TGTGCTCGATGCTTACACATGAGACAGGCGCTGTATACCGCAATGCGCTGGAGGAA 2640  
D 6625 TGTGCTCGATGCTTACACATGAGACAGGCGCTGTATACCGCAATGCGCTGGAGGAA 6684  
Q 2641 AGCAGTGGCCATCAACGCGTTGAGCAACTCTTGTGCTGCTGCTACCAACATGCTTAC 2700  
D 6685 ACCAAGCTGCCATCAATGATCACTGACCAACTCTTGTGCTGCTGCTACCAACATGCTTAC 6744  
Q 2701 GCTACCAATCCCGACCGCAGGCGCAGGCGCAGAGAGAGTCACTTTGACAGACTGCA 2760  
D 6745 GCTACCAATCCCGACCGCAGGCGCAGGCGCAGAGAGAGTCACTTTGACAGACTGCA 6804  
Q 2761 ATCTGGAAGATCACTACAGAGAGTGTCTCAAGAGATGAGAGGCGGCTCAAGTT 2820  
D 6805 ATCTGGAAGATCACTACAGAGAGTGTCTCAAGAGATGAGAGGCGGCTCAAGTT 6864  
Q 2821 AAGGCTAAGCTTATCAGTGAAGAGAGGCTGCAAGCTGACCGCCCATCTTGGGCAAA 2880  
D 6865 AAGGCTAAGCTTATCAGTGAAGAGAGGCTGCAAGCTGACCGCCCATCTTGGGCAAA 6924  
Q 2881 TCTAATTTGGCTATGAGGCGCAAGAGAGTCCGGAACCTTACAGAGGCGGTTAACAC 2940  
D 6925 TCTAATTTGGCTATGAGGCGCAAGAGAGTCCGGAACCTTACAGAGGCGGTTAACAC 6984  
Q 2941 ATCCGCTCCGTGTGAGAGACTTGTGGAAGACATGAAACCAATTGACACCAATC 3000  
D 6985 ATCCGCTCCGTGTGAGAGACTTGTGGAAGACATGAAACCAATTGACACCAATC 7044  
Q 3001 ATGGAAGAAATGAGGTTTTTTCGCTCCAAACCAAGAGAGGCGGCAAGCGCTGCG 3060  
D 7045 ATGGAAGAAATGAGGTTTTTTCGCTCCAAACCAAGAGAGGCGGCAAGCGCTGCG 7104  
Q 3061 CTTATCGTGTCCCAAGCTTGGGAGGCTCCGCTGTGAGAGAAATGAGGCTTGTGAGAGT 3120  
D 7105 CTTATCGTGTCCCAAGCTTGGGAGGCTCCGCTGTGAGAGAAATGAGGCTTGTGAGAGT 7164  
Q 3121 GTCTCAACCTCCCTCAGGCTGTGATGAGCTCTCTGTAGAGATTCAGATTTCTCTGGA 3180  
D 7165 GTCTCAACCTCCCTCAGGCTGTGATGAGCTCTCTGTAGAGATTCAGATTTCTCTGGA 7224  
Q 3181 CAGCGGCTGAGTCTCTGCTGAAACGCTTGAATCAAAAGAACCTTATGGGCTTTGCA 3240  
D 7225 CAGCGGCTGAGTCTCTGCTGAAACGCTTGAATCAAAAGAACCTTATGGGCTTTGCA 7284  
Q 3241 TATGACCGCGCTGTTTGAACATGACATGAGATGATCGGTGTGAGAGAGTCA 3300  
D 7285 TATGACCGCGCTGTTTGAACATGACATGAGATGATCGGTGTGAGAGAGTCA 7344  
Q 3301 ATTATCAATGTTGTGACTTGGCGCCGAGGCAAGCAAGGCTATGCTGCTCAGAG 3360  
D 7345 ATTATCAATGTTGTGACTTGGCGCCGAGGCAAGCAAGGCTATGCTGCTCAGAG 7404



QY 3361 CGGCTTATATCGGGGGTCCCTGACTAATTCAAAAAGGAGAACTGCGGCTATCGCGG 3420  
| | | | |  
Db 7405 CGGCTTATACATCGGGGGCCCTCGACTAATTCAAAAGGAGAACTGCGGCTATCGCGG 7464  
| | | | |  
QY 3421 TGGCGGCGGAGCGGCGCTGCTGACGACTAGTGGGCTAATACCTCACATTTACTTGAAG 3480  
| | | | |  
Db 7465 TGGCGGCGGAGCGGCTGCTGACGACCAAGCTGCGGCTAATACCTCACATTTACTTGAAG 7524  
| | | | |  
QY 3481 GCGCTGACGCTGTGAGCTGCAAAAGCTCAGAGACTGACGATGCTCGTGGGAGAG 3540  
| | | | |  
Db 7525 GCGGCTGCGGCTGTGAGCTGCAAAAGCTCAGAGACTGACGATGCTCGTGGGAGAG 7584  
| | | | |  
QY 3541 GCGCTGTCTCTTATCTGTGAGAGCGCGGAAACCCAGAGAGACGCGGAGCTTACGAGTC 3600  
| | | | |  
Db 7585 GACTTGTCTTATCTGTGAGAGCGCGGAAACCCAGAGAGACGCGGAGCTTACGAGTC 7644  
| | | | |  
QY 3601 TTCAACGAGGCTATGACTAGTACTGCTGCCCCCGGGGACCGGCCCAACGAATAC 3660  
| | | | |  
Db 7645 TTCAACGAGGCTATGACTAGTACTGCTGCCCCCGGGGACCGGCCCAACGAATAC 7704  
| | | | |  
QY 3661 GACCTGAGTTGATTAACATATGCTCTCCATGTGCGGTGCGGACGATGATCGGC 3720  
| | | | |  
Db 7705 GACTTGAAGTTGATTAACATATGCTCTCCATGTGCTCAATGTGCTAGTGGACATGCTGGC 7764  
| | | | |  
QY 3721 AAAAGGATATCTACTCACTCAACCGTGAACC 3749  
| | | | |  
Db 7765 AAAAGGATATCTACTCACTCAACCGTGAACC 7793  
| | | | |

## RESULT 6

US-10-029-907-24  
; Sequence 24, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029, 907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802)...(8407)  
US-10-029-907-24

Query Match 87.8%; Score 3291.4; DB 4; Length 8638;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;  
QY 1 TGGGAGGGGCTTTCAAGGCTTACCCAGTGGATCCCACTTCTGTCCCAACAAAG 60  
| | | | |  
Db 4046 TGGGAGGCGTCTTTACAGGCTTACCCAGTGGATCCCACTTCTGTCCCAAGACTAAG 4105  
| | | | |  
QY 61 CAGGAGGAGACAACCTTCCCTACCTGATGGCGTACAGGACTACTGTGCGGAGGGC 120  
| | | | |  
Db 4106 CAGGAGGAGACAACCTTCCCTACCTGATGGCGTACAGGACTACTGTGCGGAGGGC 4165  
| | | | |  
QY 121 CAGGCGCCACCTCCATCATGGATCAATGTGAAAGTGTCTCATACGCGCTAAAGCTTACT 180  
| | | | |  
Db 4166 CAGGCTCACTCCATCATGGATCAATGTGAAAGTGTCTCATACGCGCTAAAGCTTACT 4225  
| | | | |  
QY 181 CTGGCGGGGCGAACAACCTTGTGTATAGCTGGGAGCGCTCAAAAAGAGTCAACCTTC 240  
| | | | |  
Db 4226 CTGCACGGGCGAACAACCTTGTGTATAGCTGGGAGCGCTTCAAAAAGAGTCAACCTTC 4285  
| | | | |

QY 241 ACAACCCCATTAACCAATTCATATGAGATGATGTACCGGACTGGAGGTCGTACG 300  
| | | | |  
Db 4286 ACACACCCCATTAACCAATTCATATGAGATGATGTACCGGACTGGAGGTCGTACG 4345  
| | | | |  
QY 301 AGCACTGGGTGTGGTGGGCGGGTCTTGACAGCTCTGGCGGATTTGCTTGAACA 360  
| | | | |  
Db 4346 AGCACTGGGTGTGGTGGGCGGGTCTTGACAGCTCTGGCGGATTTGCTTGAACA 4405  
| | | | |  
QY 361 GCGAGGCTGATCATTTGGGTAGAAATCATCTTGTCCGGCGGCGGCTATTGTTCCGAC 420  
| | | | |  
Db 4406 GCGAGGCTGATCATTTGGGTAGAAATCATCTTGTCCGGCGGCGGCTATTGTTCCGAC 4465  
| | | | |  
QY 421 AAGGAAGTCTCTACAGAGATTGATGAGATGAAAGTGGCGCTGACCTCCCTTAC 480  
| | | | |  
Db 4466 AAGGAAGTCTCTTACCGGAGATTGATGAGATGAAAGTGGCGCTGACCTCCCTTAC 4525  
| | | | |  
QY 481 ATGAGACAGGAATGAGAGCTCCCGAGCGTTCAAGCAAAAAGCGTCCGGTTGCTGAG 540  
| | | | |  
Db 4526 ATGGAACAAGGAATGAGAGCTCCCGCAACAATTCMAACAAGAGCAATCGGTTGCTGAA 4585  
| | | | |  
QY 541 ACAGCACCAGAGAGCGGAGCGGCTGCTCCGCTGGTGGAGTCCAAAGTGGAGCGCTT 600  
| | | | |  
Db 4586 ACAGCACCAGAGAGCGGAGCGGCTGCTCCGCTGGTGGAGTCCAAAGTGGAGCGCTT 4645  
| | | | |  
QY 601 GAGACCTTGTGGCGGAACAATGTGAACTTCATCAGCGGATACAGTACTTGAAGAGC 660  
| | | | |  
Db 4646 GAGACCTTGTGGCGGAAGATATGTGAACTTCATCAGCGGATACCAATATTGAAGAGC 4705  
| | | | |  
QY 661 TTGTCACTCTGCTTGGGAATCCCGGATTTGATCATCTGATGGCGTTCAAGCTCTGTC 720  
| | | | |  
Db 4706 TTGTCACTCTGCTTGGGAATCCCGGATTTGATCATCTGATGGCGTTCAAGCTCTGTC 4765  
| | | | |  
QY 721 ACTAGCCGCTCAACCAACCTTCACTCTGCTTAAATCATCTGGGAGGATGGATAGCC 780  
| | | | |  
Db 4766 ACTAGCCGCTCAACCAACCTTCACTCTGCTTAAATCATCTGGGAGGATGGATAGCC 4825  
| | | | |  
QY 781 GCCCACTGCTCCCGGAGTGTGCTTCACTTGTGAGGCGCGGATGCTGTGTGCG 840  
| | | | |  
Db 4826 GCCCACTGCTCCCGGAGTGTGCTTCACTTGTGAGGCGCGGATGCTGTGTGAGGCG 4885  
| | | | |  
QY 841 GCTGTTGGAGCATAGGCTTGGGAAGTGTGTTGAGACATCTTGGCGGCTATGAGACA 900  
| | | | |  
Db 4886 GCTGTTGGAGCATAGGCTTGGGAAGTGTGTTGAGACATCTTGGCGGCTATGAGACA 4945  
| | | | |  
QY 901 GAGGTGGAGGCGCGCTGCGGCTTAAAGTATGATGAGCGGAGAAATGCCCTCACCGAG 960  
| | | | |  
Db 4946 GAGGTGGAGGCGCGCTGCGGCTTAAAGTATGATGAGCGGAGAAATGCCCTCACCGAG 5005  
| | | | |  
QY 961 GACTGTTAACTTACTTCCCTGCAATCTCTCTGTGTGCTGTGCTGTGGGATCGTG 1020  
| | | | |  
Db 5006 GACTGTTAACTTACTTCCCTGCAATCTCTCTGTGTGCTGTGCTGTGGGATCGTG 5065  
| | | | |  
QY 1021 TGGCGACGATACTGCTGTGGGACGTGGGTCCAGGGGAGGGGCTGTGAGTGAATGAAC 1080  
| | | | |  
Db 5066 TGGCGACGATACTGCTGTGGGACGTGGGTCCAGGGGAGGGGCTGTGAGTGAATGAAC 5125  
| | | | |  
QY 1081 CCGCTATAGCGTGTGCGCTCGGGGGTAAACAAGTTTCCGCCACGCACTATGAGCAAG 1140  
| | | | |  
Db 1141 AGCGACGCGACAGCATGTCTCACTCCAGATCTTCCGACTTATCTACCAACCTGTTG 1200  
| | | | |  
QY 1201 AAGAGCTTCAACGATGATTAACGAGACTGTCTCAAGCGCTGTCTCGGCTGTGGCTTA 1260  
| | | | |  
Db 5246 AAGAGCTTCAACGATGATTAACGAGACTGTCTCAAGCGCTGTCTCGGCTGTGGCTTA 5305  
| | | | |  
QY 1261 AAGGATGTTTGGAGATGATGATGACAGTTTGGCTGACCTTCAAGACTTGGCTCAAGTCC 1320  
| | | | |  
Db 5306 AAGGATGTTTGGAGATGATGATGACAGTGTGATGATTTCAAGACTTGGCTCAAGTCC 5365  
| | | | |  
QY 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATATGCCAAGTGGGTAACAAGGG 1380  
| | | | |



Db 5366 AAGCTCTGCGGATGCGGAGTCCCTTCTTCATGTCAAGTGGGTCAAGGGA 5425  
 Qy 1381 GTCTGCGGGGAGACGGCATCATGACACCTGCTCATGTGAGACAGATCACCGGA 1440  
 Db 5426 GTCTGCGGGGAGACGGCATCATGACACCTGCTCATGTGAGACAGATCACCGGA 5485  
 Qy 1441 CATGTCAAAAAGCGTTTCATGAGATCTGTGGCTTAAGACCTGTAGTAACATGTGACAT 1500  
 Db 5486 CATGTCAAAAAGCGTTTCATGAGATCTGTGGCTTAAGACCTGTAGTAACATGTGACAT 5545  
 Qy 1501 GGAACATTTCCCATCAACGCATACCAAGGCGCTCTGACGCGCTTCCCAAGCCCAAC 1560  
 Db 5546 GGAACATTTCCCATCAACGCATACCAAGGCGCTCTGACGCGCTTCCCAAGCCCAAC 5605  
 Qy 1561 TATTCAGAGGCGGTGAGGCGGGGTGCTGCTGAGAGATACGTGAGAGTTACGCGGGTGGG 1620  
 Db 5606 TATTCAGAGGCGGTGAGGCGGGGTGCTGCTGAGAGATACGTGAGAGTTACGCGGGTGGG 5665  
 Qy 1621 GATTTCACTACGTGACGACATGACCACTGACCAAGTAATAATGCGCTGCGAGTTTCA 1680  
 Db 5666 GATTTCACTACGTGACGAGCATGACCACTGACCAAGTAATAATGCGCTGCGAGTTTCCG 5725  
 Qy 1681 GCGCCCGAATTTCTTCAAGAAAGTGAATGAGGTGCGGCTGACAGGTACGCTCGGCGTGC 1740  
 Db 5726 GCGCCCGAATTTCTTCAAGAAAGTGAATGAGGTGCGGCTGACAGGTACGCTCGGCGTGC 5785  
 Qy 1741 AAAACCTCTCTACGGAGAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGTGGG 1800  
 Db 5786 AAAACCTCTCTACGGAGAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGTGGG 5845  
 Qy 1801 TCGCAGCTCCCATGCGAGCCCGAACCGGATGTAGAGTGTCACTTCCATGCTCACCGAC 1860  
 Db 5846 TCGCAGCTCCCATGCGAGCCCGAACCGGATGTAGAGTGTCACTTCCATGCTCACCGAC 5905  
 Qy 1861 CCGTCCCATCATGACAGCAGAGACGAGTAAAGCGAGCTGCGCACGAGGATCTCCCGCTCC 1920  
 Db 5906 CCGTCCCATCATGACAGCAGAGACGAGTAAAGCGAGCTGCGCACGAGGATCTCCCGCTCC 5965  
 Qy 1921 TTGGCCAGCTTTACAGTACGACATGTGTCTGGCGCTTCTCCAGAGCGCATACATTACC 1980  
 Db 5966 TTGGCCAGCTTTACAGTACGACATGTGTCTGGCGCTTCTCCAGAGCGCATACATTACC 6025  
 Qy 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCGCACTCTGAGGAGGACATGAGAG 2040  
 Db 6026 GCTCATGACTTCCCGGACGCTGACCTCATGAGGCGCACTCTGAGGAGGACATGAGAG 6085  
 Qy 2041 GCGCGGAGCATTAACCGCGTGAAGTCAAGAAACAAGTAGTAATCTGTGACTTTTCGAC 2100  
 Db 6086 GCGCGGAGCATTAACCGCGTGAAGTCAAGAAACAAGTAGTAATTTTGAATCTTTTCGAG 6145  
 Qy 2101 CCGCTCCGAGCGGAGAGATAGCGGGAAGTGTCCGTCGCGGAGATCTCTGCGGAAA 2160  
 Db 6146 CCGCTCCGAGCGGAGAGATAGCGGGAAGTGTCCGTCGCGGAGATCTCTGCGGAGG 6205  
 Qy 2161 TCCAGAAATTTCCCAACAGCGATGCCCGGATGGGACGCGCGGATTAACAACCTCGCGTGG 2220  
 Db 6206 TCCAGAAATTTCCCAACAGCGATGCCCGGATGGGACGCGCGGATTAACAACCTCGCGTGG 6265  
 Qy 2221 CTGAGTCTGAGAGGCGCGGACTACGTCCTTCCAGTGTACATGGGTGCCACTGCCA 2280  
 Db 6266 TTAGAGTCTGAGAGGCGCGGACTACGTCCTTCCAGTGTACATGGGTGTCTCATTTGGCGG 6325  
 Qy 2281 CCTACTAGAACCCCTCTCTATACCACTCCACGAGAGAGAGACAGTTGTCTGACAGAA 2340  
 Db 6326 CCTGCAAGAGCCCTCTCCGATACCACTCCACGAGAGAGAGACAGTTGTCTCTGACAGAA 6385  
 Qy 2341 TCCACCGTGTCTTTGCGCCTGCGGAGGCTTGGCCAAAGCTTTGGGTAGCTCCGACCG 2400  
 Db 6386 TCTACCGTGTCTTTGCGCCTGCGGAGGCTTGGCCAAAGCTTTGGGTAGCTCCGACCG 6445  
 Qy 2401 TCGGCGGTGACAGCGGACAGGCAACCGGCTCTGACCAACCTCCGACGAGCGGAG 2460

Db 6446 TCGGCGGTGACAGCGGACAGGCAACGCGCTCTCTGACAGCGCTCCGACGAGCGGAC 6505  
 Qy 2461 GCAGATCTGACGTTGAGTGTATTTCTCTCATATGCCCCCTTTGAGGGGAGAGCGGAGGAC 2520  
 Db 6506 GCGGAGTCCGAGGTTGAGTGTATTTCTCTCATATGCCCCCTTTGAGGGGAGAGCGGAGGANT 6565  
 Qy 2521 CCGGATCTGAGGACGCGGATCTTTGTCTACCGTGAAGTGAAGGCGCGGTGAGGACGTCTG 2580  
 Db 6566 CCGGATCTGAGGACGCGGATCTTTGTCTACCGTGAAGTGAAGGCGCGGTGAGGACGTCTG 6625  
 Qy 2581 TGTGTCTGATGTCTCTACATGAGACAGGCGCTGTATACGCGCATGCGCTGCGAGGAA 2640  
 Db 6626 TGTGTCTGATGTCTCTACATGAGACAGGCGCTGTATACGCGCATGCGCTGCGAGGAA 6685  
 Qy 2641 AGCAAGCTGCCCATCAAGCGTTGAGCACTCTTTGCTGCGGTACACAGCAACATGCTCTAC 2700  
 Db 6686 AGCAAGCTGCCCATCAAGCTGAGCACTCTTTGCTGCGGTACACAGCAACATGCTCTAC 6745  
 Qy 2701 GCTACCAATCCCGGACGCGCAAGCCGAGAGAGAGTCACTTTGACAGACTGCA 2760  
 Db 6746 GCTACCAATCTCCGAGCGGACGCGGAGAGAGTCACTTTGACAGACTGCA 6805  
 Qy 2761 ATCTGAGACGATCACTACAGAGACGTCTCAAGAGATGAAGGCGAGCGTCCACAGTT 2820  
 Db 6806 ATCTGAGACGATCACTACAGAGACGTCTCAAGAGATGAAGGCGAGCGTCCACAGTT 6865  
 Qy 2821 AAGGCTAAGCTTCTATCAATAGAGAGCTGCGCAAGCTGACGCGCCCACTTCCGCGCAA 2880  
 Db 6866 AAGGCTAAGCTTCTATCAATAGAGAGCTGCGCAAGCTGACGCGCCCACTTCCGCGCAA 6925  
 Qy 2881 TCTAAATTTGGCTATGAGGAGCAAGGACGTCGGAACCTATCCAGCAAGGCAATTAAACAC 2940  
 Db 6926 TCTAAATTTGGCTATGAGGAGCAAGGACGTCGGAACCTATCCAGCAAGGCGTTAAACAC 6985  
 Qy 2941 ATCCGCTCGGTGAGAGGACTTGTGAGAGCACTGAAACAATTTGACACACATC 3000  
 Db 6986 ATCCGCTCGGTGAGAGGACTTGTGAGAGCACTGAAACAATTTGACACACATC 7045  
 Qy 3001 ATGGCAAAAAATGAGGTTTCTGCGTCCACAGAGAGAGAGGCGCGCAAGCAGTCCG 3060  
 Db 7046 ATGGCAAAAAATGAGGTTTCTGCGTCCACAGAGAGAGAGGCGCGCAAGCAGTCCG 7105  
 Qy 3061 CTTATCGTGTCCAGACTTGGGAGGTCGCTGTGTCGAGAAATGAGCCCTCATGACG 3120  
 Db 7106 CTTATCGTGTCCAGACTTGGGAGGTCGCTGTGTCGAGAAATGAGCCCTCATGACG 7165  
 Qy 3121 GTCTCACCGTCCCTCAGGCTGTGATGGGCTCTCTGTAACGATTCAGATTTCTCTGGA 3180  
 Db 7166 GTCTCACCGTCCCTCAGGCTGTGATGGGCTCTCTGTAACGATTCAGATTTCTCTGGA 7225  
 Qy 3181 CAGCGGCTGAGTCTCTGAGTGAACGCTTGAATAAAGAGAGACCCCTATGAGGCTTTTGA 3240  
 Db 7226 CAGCGGCTGAGTCTCTGAGTGAACGCTTGAATAAAGAGAGATGCTTATGAGGCTTTTGA 7285  
 Qy 3241 TATGACACCGCTGTTTGAACCAACAGTCACTGAGATGATCATCGGTGAGAGATCA 3300  
 Db 7286 TATGACACCGCTGTTTGAACCAACAGTCACTGAGATGATCATCGGTGAGAGATCA 7345  
 Qy 3301 ATTATCAATGTTGACTTGGCCCCCGAAGCAGACAGGCGCATAGGTGCTGACAGAG 3360  
 Db 7346 ATTATCAATGTTGACTTGGCCCCCGAAGCAGACAGGCGCATAGGTGCTGACAGAG 7405  
 Qy 3361 CCGCTTTATATGAGGAGTCCCTGACTAATTAAGGAGAGAGCTGCGGCTATGCGCGG 3420  
 Db 7406 CCGCTTTATATGAGGAGTCCCTGACTAATTAAGGAGAGAGCTGCGGCTATGCGCGG 7465  
 Qy 3421 TGGCGGAGAGCGGTGCTGACGACTGACGCTGAGTATACCTTCAACATGTTTACCTTGAAG 3480  
 Db 7466 TGGCGGAGAGCGGTGCTGACGACTGACGCTGAGTATACCTTCAACATGTTTACCTTGAAG 7525  
 Qy 3481 GCGCTGAGAGCTGTCAGAGTCAAGAGCTCAAGAGCTCAAGAGCTGTCGAGAGAC 3540  
 Db 7526 GCGGCTGAGAGCTGTCAGAGTCAAGAGCTCAAGAGCTCAAGAGCTGTCGAGAGAC 7585

QY 3541 GGCCTTGTCTATCTGTGAGAGCGGGAAACCCAGAGAGACGCGGCAAGCTTACGAGTC 3600  
DB 7586 GACCTTGTCTATCTGTGAAAGCGGGGAGCCAGAGAGACGAGCGCAAGCTTACGAGTC 7645  
QY 3601 TTCACGAGAGCTTGTACTGTAGTACTGTGCCCCCGGGGACCCCGCCCAAGAAATAC 3660  
DB 7646 TTCACGAGAGCTTGTACTGTAGTACTGTGCCCCCGGGGACCCCGCCCAAGAAATAC 7705  
QY 3661 GACCTGAGTGTATATCATCATGCTCCTCCAAATGTGTGCGGTGCGGACGATGATCTGAG 3720  
DB 7706 GACCTGAGTGTATATCATCATGCTCCTCCAAATGTGTGCGGTGCGGACGATGATCTGAG 7765  
QY 3721 AAAAGGGTATACTACTCAACCCGTCGACC 3749  
DB 7766 AAAAGGGTATACTACTCAACCCGTCGACC 7794

RESULT 7  
US-10-029-907-1  
; Sequence 1, Application us/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: us/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 8639  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1803)...(8408)  
US-10-029-907-1

Query Match 87.8%; Score 3291.4; DB 4; Length 8639;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;  
QY 1 TGGGAGGGGCTCTTCAAGGCTTCAACCCAGTGAAGTCCACTTCTGTCCTCAAGGAG 60  
DB 4047 TGGGAGGGGCTCTTCAAGGCTTCAACCCAGTGAAGTCCACTTCTGTCCTCAAGGAG 4106  
QY 61 CAGGAGAGAGACAATTCCCTACCTGTGTGAGCTACAGGCTACTGTGTGAGGAGGCT 120  
DB 4107 CAGGAGAGAGACAATTCCCTACCTGTGTGAGCTACAGGCTACTGTGTGAGGAGGCT 4166  
QY 121 CAGGAGAGAGACAATTCCCTACCTGTGTGAGCTACAGGCTACTGTGTGAGGAGGCT 180  
DB 4167 CAGGCTCAACCTCATCTGTGTGAGCAAAATGTGAGAGTGTCTATACGCTTAAAGGCTAAG 4226  
QY 181 CTGCGGGGGGCAACACCTTGTGTGTATAGCTGTGAGAGGCTCCAAAGAGAGTCACTTC 240  
DB 4227 CTGCGAGGGGCAACGCGCTGTGTGTATAGGCTGTGAGGCGCTTCAAAACGAGGTTACTTAC 4286  
QY 241 ACACACCCCAATACCAATTCATCATGTGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 4287 ACACACCCCAATACCAATTCATCATGTGATGATGATGATGATGATGATGATGATGATGAT 4346  
QY 301 AGCAGCTGT 360  
DB 4347 AGCAGCTGT 4406  
QY 361 GGCAGGCTGT 420  
DB 4407 GGCAGGCTGT 4466

QY 4421 AGGGAAGTCTCTTACAGAGAGTTGATGATGAGAAAGTGTGCGCTGCGACCTTCTCTTAC 480  
DB 4467 AGGGAAGTCTCTTACAGAGAGTTGATGATGAGAAAGTGTGCGCTTCAACCTTCTCTTAC 4526  
QY 481 ATTCAGACAGGGAATGCACTGTGCGGAGAGTTCAGACAAAAGCGCTGCGGTGTCTGAG 540  
DB 4527 ATTCAGACAGGGAATGCACTGTGCGGAGAGTTCAGACAAAAGCGGTGTCTGAG 4586  
QY 541 ACAGCACCAACCAAGCGGAGGCGCTGCTCCGTTGTGTGAGTGTGCAAGTGTGAGGAGCTTC 600  
DB 4587 ACAGCACCAACCAAGCGGAGGCGCTGCTCCGTTGTGTGAGTGTGCAAGTGTGAGGAGCTTC 4646  
QY 601 GAGCCTTCTGTGAGGAGAAACATGTGAACTTCATCAGCGGATACATGACTTACGAGGCT 660  
DB 4647 GAGCCTTCTGTGAGGAGAAACATGTGAACTTCATCAGCGGATACATGACTTACGAGGCT 4706  
QY 661 TTGTCACTCTGCTGTGAGAAATCCCGGATGTGATCATATGAGCGTTCAACGCTCTGTC 720  
DB 4707 TTGTCACTCTGCTGTGAGAAATCCCGGATGTGATCATATGAGCGTTCAACGCTCTATC 4766  
QY 721 ACTAGCGGCTGACCAACCCATCTACCTCTGCTTAAATCATGCTGTGAGGAGATGAGGAGCC 780  
DB 4767 ACTAGCGGCTGACCAACCCATCTACCTCTGCTTAAATCATGCTGTGAGGAGATGAGGAGCC 4826  
QY 781 GCCCACTGTGCTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 840  
DB 4827 GCCCACTGTGCTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4886  
QY 841 GCTGTGTGAGAGATAGGCTTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 900  
DB 4887 GCTGTGTGAGAGATAGGCTTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 4946  
QY 901 GAGAGTGTGAGAGGCGGCTGTGAGCTTAAAGTATGATGAGGAGGAGAAATGCTTCCACGAG 960  
DB 4947 GAGAGTGTGAGAGGCGGCTGTGAGCTTAAAGTATGATGAGGAGGAGAAATGCTTCCACGAG 5006  
QY 961 GACCTGTGTAACTTACTTCCCTGCAATCTCTCTGTGTGAGGCTGTGTGAGGAGTGTGAG 1020  
DB 5007 GACCTGTGTAACTTACTTCCCTGCAATCTCTCTGTGTGAGGCTGTGTGAGGAGTGTGAG 5066  
QY 1021 TGGCAGAGCATACTGTGTGTGAGCACTGTGTGAGCAAGGAGGAGGAGGAGTGTGAGAGTGAAC 1080  
DB 5067 TGGCAGAGCATACTGTGTGTGAGCACTGTGTGAGCAAGGAGGAGGAGGAGTGTGAGAGTGAAC 5126  
QY 1081 CGGCTGATAGCGTGTGCTGTGAGGAGTGAACAGTGTGTGAGGAGGAGGAGGAGGAGGAGGAG 1140  
DB 5127 CGGCTGATAGCGTGTGCTGTGAGGAGTGAACAGTGTGTGAGGAGGAGGAGGAGGAGGAGGAG 5186  
QY 1141 AGGAGCGGAGAGAGTGTGATCTGATCTCTGCACTTCTGCACTTACTTCAACCACTGTG 1200  
DB 5187 AGGAGCGGAGAGAGTGTGATCTGATCTCTGCACTTCTGCACTTACTTCAACCACTGTG 5246  
QY 1201 AAGAGGCTTCAACGATGATTAACGAGAGTGTGATCTGATCTGATCTGATCTGATCTGATCT 1260  
DB 5247 AAGAGGCTTCAACGATGATTAACGAGAGTGTGATCTGATCTGATCTGATCTGATCTGATCT 5306  
QY 1261 AAGAGTGTGAGAGTGTGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1320  
DB 5307 AAGAGTGTGAGAGTGTGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 5366  
QY 1321 AAGCTCTGCGCGATTAACGAGAGTGTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1380  
DB 5367 AAGCTCTGCGCGATTAACGAGAGTGTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 5426  
QY 1381 GTCTGCGGGGAGAGCGGATCATGAGAGCACTGTCTATGTGTGAGAGAGATCACCGGA 1440  
DB 5427 GTCTGCGGGGAGAGCGGATCATGAGAGCACTGTCTATGTGTGAGAGAGATCACCGGA 5486  
QY 1441 CATGTCAAAAAGGTTCCATGTGAGAGTGTGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGT 1500  
DB 5487 CATGTCAAAAAGGTTCCATGTGAGAGTGTGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGT 5546

QY 1501 GGAACATTCCTCCATCAACGATACACCAACGAGCCCTCGACGCCCTCCAGAGGCCAAC 1560  
 DB 5547 GGAACATTCCTCCATCAACGATACACCAACGAGCCCTCGACGCCCTCCAGAGGCCAAC 5606  
 QY 1561 TATTCAGAGGCGCTGTGGCGGGTGGCTGTGAGAGATGCTGAGATTACCGAGGTGGG 1620  
 DB 5607 TATTCAGAGGCGCTGTGGCGGGTGGCTGTGAGAGATGCTGAGATTACCGAGGTGGG 5666  
 QY 1621 GATTCACATGAGTGAAGCATGACACCATGACAACTAAATGCCCGTCCAGGTCCCA 1680  
 DB 5667 GATTCACATGAGTGAAGCATGACACCATGACAACTAAATGCCCGTCCAGGTCCCA 5726  
 QY 1681 GCCCCGGAATTCCTCAACAGAGTGAATGGGGTGGCTGCAAGTACGCTCCGCGTCG 1740  
 DB 5727 GCCCCGGAATTCCTCAACAGAGTGAATGGGGTGGCTGCAAGTACGCTCCGCGTCG 5786  
 QY 1741 AAACTCTCTCTCAAGGAGAGGTCAATTCAGGTGGGCTCAACCAATACCTGGTTGG 1800  
 DB 5787 AAACTCTCTCTCAAGGAGAGGTCAATTCAGGTGGGCTCAACCAATACCTGGTTGG 5846  
 QY 1801 TGCGAGCTCCCATGCGAGCCCGGAAACCGGATGATGAGCTGCTCACTTCAGCTCACCGAC 1860  
 DB 5847 TCACAGCTCCCATGCGAGCCCGGAAACCGGATGATGAGCTGCTCACTTCAGCTCACCGAC 5906  
 QY 1861 CCTCCCATCATCAGCAGAGAGCGCTAAGCGAGGCTGCGAGGGGTCTCCCTCCCTCC 1920  
 DB 5907 CCTCCCATCATCAGCAGAGAGCGCTAAGCGAGGCTGCGAGGGGTCTCCCTCCCTCC 5966  
 QY 1921 TTGGCAGGCTCTTCACTAGCCAGTGTCTGCGCTTCTCGAAGGCGACATACATTAAC 1980  
 DB 5967 TTGGCAGGCTCTTCACTAGCCAGTGTCTGCGCTTCTCGAAGGCGACATACATTAAC 6026  
 QY 1981 CAAAATGATCTCCCAACGCTGACCTCATGAGGCCAACTCTGTGGCGCGCATGAGATG 2040  
 DB 6027 CGTCAATGATCTCCCAACGCTGACCTCATGAGGCCAACTCTGTGGCGCGCATGAGATG 6086  
 QY 2041 GGGCGGAGCATTCACCGCGTGAAGTCAAGAGAAAGATTAATCTGAGCTCTTCGAC 2100  
 DB 6087 GGGCGGAGCATTCACCGCGTGAAGTCAAGAGAAAGATTAATCTGAGCTCTTCGAC 6146  
 QY 2101 CGGCTCGAGCGGAGAGATGAGCGGAAAGTGTCCGTCGCGCGAGATCTGCGGAAA 2160  
 DB 6147 CGGCTCGAGCGGAGAGATGAGCGGAAAGTGTCCGTCGCGCGAGATCTGCGGAAA 6206  
 QY 2161 TCCAGAAATTCCTCAACGAGGATGCGCGCATGCGGACGCGGATTAACAACCTCCGCTG 2220  
 DB 6207 TCCAGAAATTCCTCAACGAGGATGCGCGCATGCGGACGCGGATTAACAACCTCCGCTG 6266  
 QY 2221 CTGAGAGCTCGGAAAGCGCGGACTGATCCCTCAATGGTATAGTGGTCCCATGCGCA 2280  
 DB 6267 CTGAGAGCTCGGAAAGCGCGGACTGATCCCTCAATGGTATAGTGGTCCCATGCGCA 6326  
 QY 2281 CCTACTAGAACCCCTCTATACCACTCCACGAGGAGAGAGACAGTTGTTCTGACAGAA 2340  
 DB 6327 CCTACTAGAACCCCTCTATACCACTCCACGAGGAGAGAGACAGTTGTTCTGACAGAA 6386  
 QY 2341 TCCAGCGTGTCTTGTGCGCTGCGGAGCTTCCCAAAAGGCTTTCCGTAAGTCCGAAACG 2400  
 DB 6387 TCCAGCGTGTCTTGTGCGCTGCGGAGCTTCCCAAAAGGCTTTCCGTAAGTCCGAAACG 6446  
 QY 2401 TGCGCGTGAAGCGGAGCGGCAACGCGGCTCTGACCAACCTCCGACGAGCGGAG 2460  
 DB 6447 TGCGCGTGAAGCGGAGCGGCAACGCGGCTCTGACCAACCTCCGACGAGCGGAG 6506  
 QY 2461 GAGAGATCTGACGTTGATGTTCTCTCATGCCCCCTTGAGGAGGAGCGGAGGAG 2520  
 DB 6507 GAGAGATCTGACGTTGATGTTCTCTCATGCCCCCTTGAGGAGGAGCGGAGGAG 6566  
 QY 2521 CCGGATCTCAGCGAGCGGCTTGGTCTACCGTGAAGGAGCGGCTGAGGAGCGTGC 2580  
 DB 6567 CCGGATCTCAGCGAGCGGCTTGGTCTACCGTGAAGGAGCGGCTGAGGAGCGTGC 6626  
 QY 2581 TGCTGCTCATGTCTTACATGAGACAGGCGCTCTGATCAAGCATGCGCTGAGGAGAA 2640

DB 6627 TGCTGCTCATGTCTTACATGAGACAGGCGCCCTGATACCGCATGCGCTGAGGAGAA 6686  
 QY 2641 AGCAAGCTGCCATCAACGCTTGAAGCACTCTTGTGCTGATACCAACAATGCTTAC 2700  
 DB 6687 ACCAAGCTGCCATCAACGCTTGAAGCACTCTTGTGCTGATACCAACAATGCTTAC 6746  
 QY 2701 GCTACCAATCCCGAGCGGAGCGGAGCGGAGAGAGGTCCTTTGACAGATGCA 2760  
 DB 6747 GCTACCAATCCCGAGCGGAGCGGAGCGGAGAGAGGTCCTTTGACAGATGCA 6806  
 QY 2761 ATCTGAGAGATCACTACAGAGAGCTGCTCAAGAGATGAGAGGCGAGGCTCACAGTT 2820  
 DB 6807 GTCCTGAGAGACACTACCGGAGCGTGTCAAGAGATGAGAGGCGAGGCTCACAGTT 6866  
 QY 2821 AAGCTAAGCTTCTATCAGTGAAGAGAGCTGCAAGCTGACGCCCCCATTCCTGGC 2880  
 DB 6867 AAGCTAAGCTTCTATCAGTGAAGAGAGCTGCAAGCTGACGCCCCCATTCCTGGC 6926  
 QY 2881 TCTAAATTTGGCTATGAGGAGGAGAGGAGGCTCCGGAACCTATCCAGAAAGCATTAAC 2940  
 DB 6927 TCTAAATTTGGCTATGAGGAGGAGAGGAGGCTCCGGAACCTATCCAGAAAGCATTAAC 6986  
 QY 2941 ATCCGCTCGTGTGGAGAGACTTGTGGAAGACACTGAAACCAATTTGACACCAATC 3000  
 DB 6987 ATCCGCTCGTGTGGAGAGACTTGTGGAAGACACTGAAACCAATTTGACACCAATC 7046  
 QY 3001 ATGGCAAAAATGAGGTTTTCTGCGTCAACAGAGAGAGAGGCGGACCAAGCTCCG 3060  
 DB 7047 ATGGCAAAAATGAGGTTTTCTGCGTCAACAGAGAGAGAGGCGGACCAAGCTCCG 7106  
 QY 3061 CTATGATGTCCCAACTTGGGGGTCGATGTGGAGAAAATGGCCCTATGAGCGT 3120  
 DB 7107 CTATGATGTCCCAACTTGGGGGTCGATGTGGAGAAAATGGCCCTATGAGCGT 7166  
 QY 3121 GTCTCACCTCCCTCAGGCTGATGAGGCTCCTCGTACGAAATTCAGATTTCTCCGTA 3180  
 DB 7167 GTCTCACCTCCCTCAGGCTGATGAGGCTCCTCGTACGAAATTCAGATTTCTCCGTA 7226  
 QY 3181 CAGCGGCTGAGTTCCTGTGAAACGCTGGAATCAAAGAGACCCCTATGAGCTTGA 3240  
 DB 7227 CAGCGGCTGAGTTCCTGTGAAATGCTGGAATCAAAGAGACCCCTATGAGCTTGA 7286  
 QY 3241 TATGACACCGGCTTTTGAACAGTCACTGAAAGTGAATGCAATCCGTATGAGAGTCA 3300  
 DB 7287 TATGACACCGGCTTTTGAACAGTCACTGAAAGTGAATGCAATCCGTATGAGAGTCA 7346  
 QY 3301 ATTATCAATGTGTGACTTGGCCCCCGAGAGCGACAGGCTATAGTCTCACAGAG 3360  
 DB 7347 ATTATCAATGTGTGACTTGGCCCCCGAGAGCGACAGGCTATAGTCTCACAGAG 7406  
 QY 3361 CGGCTTATATCGGAGGCTCCCTGACTAATCAAAAAGGAGAGTGGCGCTATGCGCG 3420  
 DB 7407 CGGCTTATATCGGAGGCTCCCTGACTAATCAAAAAGGAGAGTGGCGCTATGCGCG 7466  
 QY 3421 TGGCGGAGAGGCGTGTGAGAGTACTAGTGGGTAATACCTTCAATGTTACTTGAAG 3480  
 DB 7467 TGGCGGAGAGGCGTGTGAGAGTACTAGTGGGTAATACCTTCAATGTTACTTGAAG 7526  
 QY 3481 GCCTGAGAGCTGAGAGCTCAAAAGCTCAGAGATGACAGTGTCTGTGGAGAG 3540  
 DB 7527 GCCTGAGAGCTGAGAGCTCAAAAGCTCAGAGATGACAGTGTCTGTGGAGAG 7586  
 QY 3541 GGCCTTGTGCTATCTGAGAGCGGAGAACCCAGAGAGAGCGGCGAGCTTACGAGT 3600  
 DB 7587 GGCCTTGTGCTATCTGAGAGCGGAGAACCCAGAGAGAGCGGCGAGCTTACGAGT 7646  
 QY 3601 TTCAAGGAGGCTATGACTGATGATCTTGGCCCCCGGAGAGCCCGCCCAACAAGATAC 3660  
 DB 7647 TTCAAGGAGGCTATGACTGATGATCTTGGCCCCCGGAGAGCCCGCCCAACAAGATAC 7706  
 QY 3661 GACCTGAGATTAATCATGATCTCTCAATGTTGCGTGGCGACGATCATCTGCG 3720

Db 7707 GACTTGAGTTGATTAACATCATGCTCTCCATGATGTGCTGACGCGGACGATGATCTGGC 7766  
QY 3721 AAAAGGATATCTACTACCTCACCCCGTGACC 3749  
Db 7767 AAAAGGATATCTACTACCTCACCCCGTGACC 7795

RESULT 8  
US-09-539-601-13  
Sequence 13, Application US/09539601C  
Patent No. 6630343  
GENERAL INFORMATION:  
APPLICANT: Bartschlagel, Ralf FW  
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
FILE REFERENCE: all sequences  
CURRENT APPLICATION NUMBER: US/09/539,601C  
EARLIER APPLICATION NUMBER: 2001-08-30  
EARLIER FILING DATE: 1999-04-03  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 13  
LENGTH: 8649  
TYPE: DNA  
ORGANISM: Hepatitis C virus  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: (1) .. (341)  
OTHER INFORMATION: construct I389/NS2-3'/wt  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (342) .. (1193)  
OTHER INFORMATION: hepatitis c virus core-neomycin phosphotransferase  
OTHER INFORMATION: fusion protein  
FEATURE:  
NAME/KEY: RBS  
LOCATION: (1202) .. (1812)  
OTHER INFORMATION: internal ribosome entry site of  
OTHER INFORMATION: encephalomyocarditis virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1813) .. (8418)  
OTHER INFORMATION: hepatitis C virus NS2 - 5B  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: (8419) .. (8649)  
PUBLICATION INFORMATION:  
AUTHORS: Lohmann, Volker  
AUTHORS: Koch, Jan-Oliver  
AUTHORS: Herian, Ulrike  
AUTHORS: Theilmann, Lorenz  
AUTHORS: Bartschlagel, Ralf  
TITLE: Replication of subgenomic hepatitis c virus RNAs in a  
JOURNAL: Science  
VOLUME: 285  
PAGES: 110-113  
DATE: 1999-07-02  
US-09-539-601-13

Query Match 87.8%; Score 3291.4; DB 4; Length 8649;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3463; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 1 TGGAGAGGCGCTTTTACAGGCTTACCCACGTGAGTCCCACTTCTGTCCTCAACCAAG 60  
Db 4057 TGGAGAGGCGCTTTTACAGGCTTACCCACCATAGAGCCCATTTCTGTCCAGACTAAG 1116  
QY 61 CAGGAGGAGACACTTCCCTACCTGAGCGTACCGAGCTACTGTGTCGCTGAGGCC 120  
Db 4117 CAGGAGGAGACACTTCCCTACCTGAGCGTACTACGAGCTACGAGTGCAGGAGCT 4176

QY 121 CAGGCCCACTTCATCATGAGATCAATGTGAAAGTGTCTCAATACGCTAAAGCTTACT 180  
Db 4177 CAGGCTTCACCTTCATCGTGGACCAATGTGAAAGTGTCTCAATACGCTAAAGCTTACT 4236  
QY 181 CTGCGCGGGCCCAACCCCTTGTGTATAGGCTGTGGAGCGGCTCAAAAGAGGTACCCGTC 240  
Db 4237 CTGCAAGGGCCAAACGCCCTGTGTATAGGCTGTGGAGCGGCTCAAAAGAGGTATACCC 4296  
QY 241 ACACACCCCATTAACCAATTTCAATGAGCATGTCAGCCGACCTGAGAGTGTGTACAG 300  
Db 4297 ACACACCCCATTAACCAATTAATCATGAGCATGTCAGCTGTGAGAGGTCGTACAG 4356  
QY 301 AGCACTGGGTGTGTGTGGCGGGGTCTTGTGAGCTGTGGCTGTGATTTGCTTGAACA 360  
Db 4357 AGCACTGGGTGTGTGTGGCGGGGTCTTGTGAGCTGTGGCTGTGATTTGCTTGAACA 4416  
QY 361 GGCAGGTGTGTATGT 420  
Db 4417 GGCAGGTGTGTATGT 4476  
QY 421 AGGAAAGTCTCTACAGAGATTCAGATGAGATGAAAGTGTGTGTGTGTGTGTGTGTGT 480  
Db 4477 AGGAAAGTCTCTTACCGGAGATTCAGATGAGATGAAAGTGTGTGTGTGTGTGTGTGT 4536  
QY 481 ATCGAGCAGGAAATGACGCTGCGCGAGCATTCAGCAAAAAGCGTGGGTGTGTGAC 540  
Db 4537 ATCGAACAAGGAATGACGCTGCGCGAACAATTCAGCAAAAAGCGTGGGTGTGTGAC 4596  
QY 541 ACAGCCACCAACCAAGCGGAGCGGCTGCTCCGCTGTGTGTGTGTGTGTGTGTGTGT 600  
Db 4597 ACAGCCACCAACCAAGCGGAGCGGCTGCTCCGCTGTGTGTGTGTGTGTGTGTGTGT 4656  
QY 601 GAGACCTTGTGGGCGAAACACATGTGAACTTCATCAGCGGATACAGTACTTACAGAGC 660  
Db 4657 GAAAGCTTGTGGGCGAAAGATATGTGATTTTATCAGGCGGATACATATTATAGAGGC 4716  
QY 661 TTGTCCACTGTGCTGTGGAAATCCCGGATGTGATCATGTAGCGTTCAAGCTCTGTTC 720  
Db 4717 TTGTCCACTGTGCTGTGGAAATCCCGGATGTGATCATGTAGCGATTCACAGCTCTATTC 4776  
QY 721 ACTAGCCGCTTACCAACCCATTCATCCCTGTCTTAATCTGTGGGGGATGTGTAGGC 780  
Db 4777 ACTAGCCGCTTACCAACCCATTCATCCCTGTCTTAATCTGTGGGGGATGTGTAGGC 4836  
QY 781 GCCCACTGCTCCCCCAGT 840  
Db 4837 GCCCACTGCTCTCTCCAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4896  
QY 841 GCTGTGTGAGCATAGGCTTGTGGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900  
Db 4897 GCTGTGTGAGCATAGGCTTGTGGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4956  
QY 901 GGAATGTGAGGCGGCTGT 960  
Db 4957 GGAATGTGAGGCGGCTGT 5016  
QY 961 GACCTGTTAATCTTATCTCTGTCATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020  
Db 5017 GACCTGTTAATCTTATCTCTGTCATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 5076  
QY 1021 TGGCAGCAGATATCTGT 1080  
Db 5077 TGGCAGCAGATATCTGT 5136  
QY 1081 CGGCTGATATGCTTGT 1140  
Db 5137 CGGCTGATATGCTTGT 5196  
QY 1141 AGGAGCGCGAGCAGATGT 1200  
Db 5197 AGGAGCGCGAGCAGATGT 5256  
QY 1201 AAGAGCTTCAACCATGTGATTAACGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260

Db 5257 |AAGAGCTTCACCAAGGATCAACGAGAGCTGCTCAACGCAATGCTCCGGCTGTGGCTA 5316  
 Qy 1261 |AAGGATGTTGGGAGCTGATATGACAGATTTGGCTGACTTAAAGACTTGGCTCCAGTCC 1320  
 Db 5317 |AAGAGATGTTGGGATGATATGACAGGTTGATGATTTCAAGACTTGGCTCCAGTCC 5376  
 Qy 1321 |AAGCTTCGCGCGATTAACCGGAGTCCCTTTTCTCATGCAAGTGGGTACAAGGGG 1380  
 Db 5377 |AAGCTTCGCGCGATTAACCGGAGTCCCTTTTCTCATGCAAGTGGGTACAAGGGG 5436  
 Qy 1381 |GTTCGCGGGGAGACGGCATCATGACAGACCACTGCTCATGTGGAGACAGATCACCGGA 1440  
 Db 5437 |GTTCGCGGGGAGACGGCATCATGACAGACCACTGCTCATGTGGAGACAGATCACCGGA 5496  
 Qy 1441 |CATGTCAAAAACGGTTTCATGAGGATGTTGGGCTTAAGAAGCTGTAGTAAACATGTGCAT 1500  
 Db 5497 |CATGTCAAAAACGGTTTCATGAGGATGTTGGGCTTAAGAAGCTGTAGTAAACATGTGCAT 5556  
 Qy 1501 |GGAACCTTCCCATCAACGATACACCAAGGAGCCCTTGCACGCGCTCCCAAGGCCAAG 1560  
 Db 5557 |GGAACCTTCCCATCAACGATACACCAAGGAGCCCTTGCACGCGCTCCCAAGGCCAAG 1566  
 Qy 1561 |TATTCAGGCGCTGTGGCGGCTGCTGAGAGTACGTGAGGTTACGCGGCTGGG 1620  
 Db 5617 |TATTCAGGCGCTGTGGCGGCTGCTGAGAGTACGTGAGGTTACGCGGCTGGG 5676  
 Qy 1621 |GATTTCCACTACGTGACGAGCATGACCACTGACAACTAAATGCCGTGCGAGTTCCA 1680  
 Db 5677 |GATTTCCACTACGTGACGAGCATGACCACTGACAACTAAATGCCGTGCGAGTTCCA 5736  
 Qy 1681 |GCCCCGGAATTTCTTCAAGAGTGAATGGGCTGCGCTGACAGTACGCTCCGGCTGC 1740  
 Db 5737 |GCCCCGGAATTTCTTCAAGAGTGAATGGGCTGCGCTGACAGTACGCTCCGGCTGC 5796  
 Qy 1741 |AAACCTCTCTACGGAGAGGTCACTTCAGGTGGGCTCAACCAATACCTGTGGG 1800  
 Db 5797 |AAACCTCTCTACGGAGAGGTCACTTCAGGTGGGCTCAACCAATACCTGTGGG 5856  
 Qy 1801 |TCGCAAGCTCCATGCGAGCCGCAACCGATGAGAGTGTAGAGTGTCTCATGCTCACCGAC 1860  
 Db 5857 |TCAGAGCTCCATGCGAGCCGCAACCGATGAGAGTGTAGAGTGTCTCATGCTCACCGAC 5916  
 Qy 1861 |CCCTCCCATCATCAGAGAGAGAGCGCTAAGCGAGCTGCGCAGGGGTTCCCCCTCC 1920  
 Db 5917 |CCCTCCCATCATCAGAGAGAGAGCGCTAAGCGAGCTGCGCAGGGGATCTCCCTCC 5976  
 Qy 1921 |TTGGCCAGCTTCAGCTAGCCAGTTGTCTGGGCTTCTCGAAGGCGACATCATTTAC 1980  
 Db 5977 |TTGGCCAGCTTCAGCTAGCCAGTTGTCTGGGCTTCTCGAAGGCGACATCATTTAC 6036  
 Qy 1981 |CAAAATGACTTCCACAGCTGACCTCATGAGGCGCAACCTCTGTGGCGGCAATGAGATG 2040  
 Db 6037 |CTGATGACTTCCCGGAGCGTGAACCTCATGAGGCGCAACCTCTGTGGCGGCAATGAGATG 6096  
 Qy 2041 |GGGCGGAGCATTAACCGCGTGAAGTCAAGAGACAAGTAAATCTGTGACTTTTGAC 2100  
 Db 6097 |GGGCGGAGCATTAACCGCGTGAAGTCAAGAGACAAGTAAATCTGTGACTTTTGAC 6156  
 Qy 2101 |CGGCTCCGAGCGGAGAGATGAGCGGGAAGTCCGTCGCGGAGAGATCCGCGGAA 2160  
 Db 6157 |CGGCTCCGAGCGGAGAGATGAGCGGGAAGTCCGTCGCGGAGAGATCCGCGGAA 6216  
 Qy 2161 |TCCAGAAATTTCCACAGCGATGCGCGCATGGGCAAGCCGAGTTACAAACCTCCGCTG 2220  
 Db 6217 |TCCAGAAATTTCCCGAGGAGTCCCATATGGGCAAGCCGAGTTACAAACCTCCCATG 6276  
 Qy 2221 |CTGAGATCTGGAAGGCGCGGACTAGTCTCTCAGTGTGTAATGAGGTGCCACTGCCA 2280  
 Db 6277 |TTAGATCTCTGGAAGGCGCGGACTAGTCTCTCAGTGTGTAATGAGGTGCCACTGCCG 6336  
 Qy 2281 |CCACTAAGACCCCTCTATACACCTCCAGAGGAGAGAGACAGTTGTTCTGACGAA 2340

Db 6337 |CTGCAAGGCCCTCCGATACCACTCCAGGAGGAAGAGAGGTTGCTGTGAGAA 6396  
 Qy 2341 |TCCACCGTGTCTTTCGCTGCGGAGCTTTCGCAAGAGCTTTGGTGTGCTCCGAACG 2400  
 Db 6397 |TCCACCGTGTCTTTCGCTGCGGAGCTTTCGCAAGAGCTTTGGAGTGTCCGACTCCGAA 6456  
 Qy 2401 |TCGGCCGTGACAGCGGAGCGCAACCGGCGCTCTGCAACCACTCCGACGAGCGGAG 2460  
 Db 6457 |TCGGCCGTGACAGCGGAGCGCAACCGGCGCTCTGCAACCACTCCGACGAGCGGAG 6516  
 Qy 2461 |GCAGATCTGACGTTGATGTTATTCCTCATGCCCCCTTTGAGGGGAGCCGCGGAG 2520  
 Db 6517 |GGGAGTCCGAGCGTTAAGTGTACTCTCATGCCCCCTTTGAGGGGAGCCGCGGAG 6576  
 Qy 2521 |CCGATCTCAGGAGCGGCTTGTGTACCGTGAAGTGAAGGCGGTGAAGACGTCCTC 2580  
 Db 6577 |CCGATCTCAGGAGCGGCTTGTGTACCGTGAAGTGAAGGCGGTGAAGACGTCCTC 6636  
 Qy 2581 |TGCTGCTCGATGCTTACACATGAGACAGGCGCTGATCACGCGCATGCGTGAGAGAA 2640  
 Db 6637 |TGCTGCTCGATGCTTACACATGAGACAGGCGCTGATCACGCGCATGCGTGAGAGAA 6696  
 Qy 2641 |AGCAAGCTGCCATCAACGCGTTGAGCAACTTTTGTGCTGCTACCAACATGCTTAC 2700  
 Db 6697 |ACCAAGCTGCCATCAACGCGTTGAGCAACTTTTGTGCTGCTACCAACATGCTTAC 6756  
 Qy 2701 |GCTACACATCCCGCAGCGGAGCCAGGCGGAGAGAGTCACTTTGACAGACTGCCA 2760  
 Db 6757 |GCTACACATCCCGCAGCGGAGCCAGGCGGAGAGAGTCACTTTGACAGACTGCCA 6816  
 Qy 2761 |ATCCTGGAAGATCACTACAGAGAGCGTCAAGAGATGAGAGGCGAGGCGTCCAGATT 2820  
 Db 6817 |GTCTGGAAGATCACTACAGAGAGCGTCAAGAGATGAGAGGCGAGGCGTCCAGATT 6876  
 Qy 2821 |AAGCTAAGCTTCTATCAGTAGAGAGCGCTGCAAGCTGACGCGCCCACTTGGGCAAA 2880  
 Db 6877 |AAGCTAAGCTTCTATCAGTAGAGAGCGCTGCAAGCTGACGCGCCCACTTGGGCAAA 6936  
 Qy 2881 |TCTAAATTTGGCTATGGGCAAGAGAGCTCCGGAACCTTACGCAAGGCGCATTTAAC 2940  
 Db 6937 |TCTAAATTTGGCTATGGGCAAGAGAGCTCCGGAACCTTACGCAAGGCGCATTTAAC 6996  
 Qy 2941 |ATCGGCTCGTGGGAGAGACTTGTGGAAGACACTGAACCAATTTGACACACATC 3000  
 Db 6997 |ATCGGCTCGTGGGAGAGACTTGTGGAAGACACTGAACCAATTTGACACACATC 7056  
 Qy 3001 |ATGCAAAAAATGAGGTTTCTGCTCAACAGAGAGAGAGGCGCAAGCCAGCTGCG 3060  
 Db 7057 |ATGCAAAAAATGAGGTTTCTGCTCAACAGAGAGAGAGGCGCAAGCCAGCTGCG 7116  
 Qy 3061 |CTTATGCTGTTCCAGACTTGGGAGTCCGTGTGCGAGAAATGGCCCTTATGACGTG 3120  
 Db 7117 |CTTATGCTGTTCCAGACTTGGGAGTCCGTGTGCGAGAAATGGCCCTTATGACGTG 7176  
 Qy 3121 |GTCTCAACCTCCCTAGGCTGATGAGGCTCCTGTAGAGATTCAGTATTCCTCGGA 3180  
 Db 7177 |GTCTCAACCTCCCTAGGCTGATGAGGCTCCTGTAGAGATTCAGTATTCCTCGGA 7236  
 Qy 3181 |CAGCGGCTGAGTTCCTGTGAAAGCGCTGGAATCAAGAGAGAGCCCTATGGCTTTGCA 3240  
 Db 7237 |CAGCGGCTGAGTTCCTGTGAAAGCGCTGGAATCAAGAGAGAGCCCTATGGCTTTGCA 7296  
 Qy 3241 |TATGACACCGGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTGAGAGTCA 3300  
 Db 7297 |TATGACACCGGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTGAGAGTCA 7356  
 Qy 3301 |ATTATCAATGTTGATGACTTGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
 Db 7357 |ATTATCAATGTTGATGACTTGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7416  
 Qy 3361 |CGGCTTTATATCGGAGGCTCCCTGACTAATTCAAAAGGCGAGAACTGCGGCTATCGCG 3420  
 Db 7417 |CGGCTTTATATCGGAGGCTCCCTGACTAATTCAAAAGGCGAGAACTGCGGCTATCGCG 7476



```
Db 7564 CGGCTATAGCGTTTCGCTTCGCGGGGTAAACAGTCTCCCCCAAGCATATGTGGCTGAG 7623
QY 1141 AGGAGAGCCGAGACAGTGTCACTAGATCTCTCCGACTTACTATACCCCACTGTG 1200
Db 7624 AGGAGAGCCGAGACAGTGTCACTAGATCTCTCCGACTTACTATACCCCACTGTG 7683
QY 1201 AAGAGGCTCCACAGATGATTAACAGAGAGTGTCTCCAGCGCCCTGCTCCGGCTGTGCGTA 1260
Db 7684 AAGAGGCTCCACAGATGATTAACAGAGAGTGTCTCCAGCGCCCTGCTCCGGCTGTGCGTA 7743
QY 1261 AAGAGATGTTGGAGATGATATGACAGAGTGTGCTGACTTCAAGACTGTGCTCACTCC 1320
Db 7744 AAGAGATGTTGGAGATGATATGACAGAGTGTGCTGACTTCAAGACTGTGCTCACTCC 7803
QY 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCAAGTGGGTACAAAGGG 1380
Db 7804 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCAAGTGGGTACAAAGGG 7863
QY 1381 GTCTGGCGGGAGACGAGCATATGACAGACCACTGCTCATGTGAGAGACAGATCACCGGA 1440
Db 7864 GTCTGGCGGGAGACGAGCATATGACAGACCACTGCTCATGTGAGAGACAGATCACCGGA 7923
QY 1441 CATGTCMAAAACGGTTCATGAGAGATGTTGGGCTTAAGACTGTAGTAACTATGTGCAAT 1500
Db 7924 CATGTCMAAAACGGTTCATGAGAGATGTTGGGCTTAAGACTGTAGTAACTATGTGCAAT 7983
QY 1501 GGAACATTTCCCATCAACGCAATACCAACGAGCCCTGCAAGCCCTTCCCAAGCCCAAC 1560
Db 7984 GGAACATTTCCCATCAACGCAATACCAACGAGCCCTTCCCAAGCCCTTCCCAAGCCCAAC 8043
QY 1561 TATTCAGAGGCGGTGAGGCGGGTGGCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1620
Db 8044 TATTCAGAGGCGGTGAGGCGGGTGGCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 8103
QY 1621 GATTTCCACTACGTCGAGAGATGACCACTGCAACAGTAAATGCCGCTGCAAGTGTCCA 1680
Db 8104 GATTTCCACTACGTCGAGAGATGACCACTGCAACAGTAAATGCCGCTGCAAGTGTCCA 8163
QY 1681 GCCCGGAAATTTCTTCAACAGAGTGAATGGGGTGGCGGTGCAAGTGAAGTGAAGTGAAG 1740
Db 8164 GCCCGGAAATTTCTTCAACAGAGTGAATGGGGTGGCGGTGCAAGTGAAGTGAAGTGAAG 8223
QY 1741 AAACCTCTCTACAGGAGAGATGATTCATTCAGTGGGCTCAACCAATCTCTGTGGTGG 1800
Db 8224 AAACCTCTCTACAGGAGAGATGATTCATTCAGTGGGCTCAACCAATCTCTGTGGTGG 8283
QY 1801 TCGAGACTCCATGCGAGCCGGAACGAGATGTAGCAGTGTCTCACTTCCATGCTCACCGAC 1860
Db 8284 TCGAGACTCCATGCGAGCCGGAACGAGATGTAGCAGTGTCTCACTTCCATGCTCACCGAC 8343
QY 1861 CCGTCCCAATCAACAGCAGAGACGAGCTAAGCGAGGTGGCAAGGGGGTCTCCCTCTCC 1920
Db 8344 CCGTCCCAATCAACAGCAGAGACGAGCTAAGCGAGGTGGCAAGGGGGTCTCCCTCTCC 8403
QY 1921 TTGGCAGGCTCTTCACTAGCCAGTGTCTGCGGCTTCTCGAAGCGAGATACATTAAC 1980
Db 8404 TTGGCAGGCTCTTCACTAGCCAGTGTCTGCGGCTTCTCGAAGCGAGATACATTAAC 8463
QY 1981 CAAATATCTTCCCAACGCTGACCTCATGAGGCAACCTCTGTGGCGGATGAGATG 2040
Db 8464 CCGATATCTTCCCGGAGCGGTGACCTCATGAGGCAACCTCTGTGGCGGAGAGATG 8523
QY 2041 GCGGAGGACATTAACCGCGGTGAGTCAAGAGACAAAGTAAATCTCTGACCTTTTGAC 2100
Db 8524 GCGGAGGACATTAACCGCGGTGAGTCAAGAGATTAAGTAAATTTTGGACTTTTGAC 8583
QY 2101 CCGCTCCAGCGGAGAGATGAGCGGGAATGTCCTCCCGCGGAGATCTCTCGGAAA 2160
Db 8584 CCGCTCCAGCGGAGAGATGAGCGGGAATGTCCTCCCGCGGAGATCTCTCGGAAA 8643
QY 2161 TCCAAATAATTCACACAGGATGCGGCAATGGGCAAGCCGGAATTAACAACCTCTCGCTG 2220
Db 8644 TCCAAATAATTCCTCAGAGATGCCCATATGGGCAAGCCCGGAATTAACAACCTCTCGCTG 8703
QY 2221 CTGAGATCCGGAAGGCGCCCGGACTACGTCCTCCATGAGTACATGAGTGGCCCACTGCA 2280
Db 8704 TTAGAGTCTCTGGAAGGACCGGACTACGTCCTCCATGAGTACATGAGTGGCCCACTGCA 8763
QY 2281 CTTACTAAGACCCCTCTTATACCACTCCACGAGAGAGAGACAGTGTCTGACAGAA 2340
Db 8764 CTTACTAAGACCCCTCTCAGTATACCACTCCACGAGAGAGAGAGAGTGTCTGCTGAGAA 8823
QY 2341 TCCACCGTGTCTTCTGCTGCGGAGTGTGCAAAAGCTTTCGTAAGTCTGCAACG 2400
Db 8824 TCCACCGTGTCTTCTGCTGCGGAGTGTGCAAAAGCTTTCGTAAGTCTGCAACG 8883
QY 2401 TCGGCGGTGACAGCGGCAAGGCAAGGCGCCCTCCGACCAACCTCCGACGAGCGGGA 2460
Db 8884 TCGGCGGTGACAGCGGCAAGGCAAGGCGCCCTCCGACCAAGCTCCGACGAGCGGGA 8943
QY 2461 GCAGATCTGACGCTGAGTCTGATTTCTCCATGCCCCCTTGAAGGGGAGCGGCGGAC 2520
Db 8944 GCAGATCTGACGCTGAGTCTGATTTCTCCATGCCCCCTTGAAGGGGAGCGGCGGAT 9003
QY 2521 CCGATCTCAGCGAGCGGCTTGTGTCTACCTGAGTGAAGGCGGCTGAGACGTCGTG 2580
Db 9004 CCGATCTCAGCGAGCGGCTTGTGTCTACCTGAGTGAAGGCGGCTGAGACGTCGTG 9063
QY 2581 TGTGTGTGATGCTTCAACATGAGAGAGGCGCTGATACCGCATAGGCGCTGCGGAGAA 2640
Db 9064 TGTGTGTGATGCTTCAACATGAGAGAGGCGCTGATACCGCATAGGCGCTGCGGAGAA 9123
QY 2641 AGCAAGCTCCCATCAACGCGTGTGAGCAACTCTTGTCTGCTGCTGCAACCAATGCTCTAC 2700
Db 9124 AGCAAGCTCCCATCAACGCGTGTGAGCAACTCTTGTCTGCTGCTGCAACCAATGCTCTAC 9183
QY 2701 GCTACCAATCTCCGAGCGGCAAGGCGGAGAGAGAGTCACTTTTACAGACTGCA 2760
Db 9184 GCTACCAATCTCCGAGCGGCAAGGCGGAGAGAGAGTCACTTTTACAGACTGCA 9243
QY 2761 ATCTGGAAGATCACTACAGAGACGTCCTCAAGAGATGAAGCGGACGCTCAAGT 2820
Db 9244 ATCTGGAAGATCACTACAGAGACGTCCTCAAGAGATGAAGCGGACGCTCAAGT 9303
QY 2821 AAGCTTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACGCCCCCATTCGAGCAAA 2880
Db 9304 AAGCTTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACGCCCCCATTCGAGCAAA 9363
QY 2881 TCTAAATTTGGCTATGAGGCAAGAGAGCTCGGAACTTATCAAGAGGCACTTAAACAC 2940
Db 9364 TCTAAATTTGGCTATGAGGCAAGAGAGCTCGGAACTTATCAAGAGGCACTTAAACAC 9423
QY 2941 ATCTGCTCCGTGTGAGAGAGCTTGTGAGAGACATGAAACCAATTTGACACACATC 3000
Db 9424 ATCTGCTCCGTGTGAGAGAGCTTGTGAGAGACATGAAACCAATTTGACACACATC 9483
QY 3001 ATGCAAAAATGAGGTTTTTCTGCGTCAACAGAGAGAGAGGCGGCAAGCAAGCTGCG 3060
Db 9484 ATGCAAAAATGAGGTTTTTCTGCGTCAACAGAGAGAGAGGCGGCAAGCAAGCTGCG 9543
QY 3061 CTTATGATGTTCCAGACTTGGGGGTCCGTGTGTGAGAGAAATGGCCCTTATGACGTG 3120
Db 9544 CTTATGATGTTCCAGACTTGGGGGTCCGTGTGTGAGAGAAATGGCCCTTATGACGTG 9603
QY 3121 GTCTCAACCTCCCTCAGAGCTGTGATGGGCTCTCTGATGAGATTCAGATTTCTCTGGA 3180
Db 9604 GTCTCAACCTCCCTCAGAGCTGTGATGGGCTCTCTGATGAGATTCAGATTTCTCTGGA 9663
QY 3181 CAGCGGATGAGTCTCTGATGAAAGCGCTGAAATCAAGAGACCCCTTATGAGCTTTGCA 3240
Db 9664 CAGCGGATGAGTCTCTGATGAAAGCGCTGAAATCAAGAGACCCCTTATGAGCTTTGCA 9723
QY 3241 TATGACACCGGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 3300
Db 9724 TATGACACCGGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 9783
```



QY	3301	ATTATTAACAATGTTGACTTGGCCCCCGGAAGCCAGACAGGSCATTAAGTGTCTCAGAGAG	3360
Db	9784	ATCTACCAATGTGTGACTTGGCCCCCGGAAGCCAGACAGGSCATTAAGTGTCTCAGAGAG	9843
QY	3361	CGGCTTTATATCGGGGGGTCCCTGTACATTAATTCAAAAGGAGAGACTGGCGCTTATCGCCG	3420
Db	9844	CGGCTTTATACATCGGGGGGCCCTGTACATTAATTCAAAAGGAGAGACTGGCGCTTATCGCCG	9903
QY	3421	TGCCCCCGGAGCCGGCGTGTGACGACTGACCTAGCTGCGGTAAATACCTCACAATGTTACTTTGAAG	3480
Db	9904	TGCCCCCGGAGCCGGGTGTACTGACGACAGCGCTGCGGTAAATACCTCACAATGTTACTTTGAAG	9963
QY	3481	GCGTCTGAGCGCTTCCAGCTGTCAAAAGCTCCAGAGCTGCACGATGCTCGTGTGCGAGAC	3540
Db	9964	GCGCTGTGCGGCTGTCCAGCTGTCAAAAGCTCCAGAGCTGCACGATGCTCGTGTGCGAGAC	10022
QY	3541	GGCGTTGTCTGTAATCTGTGAGAGCGCGGGAAACCCAGAGGAGCGCGGCGAGCTTACGAGTTC	3600
Db	10024	GACCTTGTCTGTAATCTGTGAGAGCGCGGGAAACCCAGAGGAGCGCGGCGAGCTTACGAGTTC	10083
QY	3601	TTTCAAGGAGGTATGACTGACTGACTCTGGCCCCCGGGGAGCCGCCGCCAACCAGAAATAC	3660
Db	10084	TTTCAAGGAGGTATGACTGACTGACTCTGGCCCCCGGGGAGCCGCCGCCAACCAGAAATAC	10143
QY	3661	GACCTGAGGTGATTAAGTACATCATGCTCTCCCAATGTGTGCGTGTGCGACGATGATCTGGC	3720
Db	10144	GACCTGAGGTGATTAAGTACATCATGCTCTCCCAATGTGTGTCAATGTGCGGACGATGATCTGGC	10203
QY	3721	AAAAAGGTATCTACCTTCACCCGTGAGCC 3749	
Db	10204	AAAAAGGTATCTATCTCACCCGTGAGCC 10232	

```

RESULT 10
US-10-029-907-6
; Sequence 6, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
US-10-029-907-6

```

Query Match	87.7%;	Score 3288.2;	DB 4;	Length 8658;
Best Local Similarity	92.3%;	Pred. No. 0;		
Matches 3461; Conservative	0;	Mismatches 288;	Indels 0;	Gaps 0;

Qy	1	TGGGAGGGCGCTTTCACAGGGCTACCCAGTGGATGCCCACTTCGTGTCCAAACAAG	60
Db	4046	TGGGAGAGCGTCTTTCACAGGCTCACCCACATACGCCCATTTCTGTGCCAGACTAAG	4105
Qy	61	CAGCGAGAGACAACCTTCCCTTACTCGTGTGGCGTACCAAGGCTACTGTGTCCGCTAAGGCC	120
Db	4106	CAGCGAGAGACAACCTTCCCTTACTCGTGTACATACCAAGGCTACGGTGTCCGCGCAGGGCT	4165
Qy	121	CAGGCCCCACTTCATCATGGGATCAATATGGAAAGTGTCTCATACGGCTAAAGCCTACT	180
Db	4166	CAGGCTCACCTTCATCGTGGAGCAAAATGGAAAGTGTCTCATACGGCTAAAGCCTACG	4225

QY	181	TTGGCGGGGGCCAAACACCCCTGTCTGTATAGGTGGAGACCGCGTCCAAACAGAGTCAACCTC	240
Db	4226	CTGACCGGGCCAAAGCCCTCTGTCTATAGGTGGAGACCGTTTCAAACAGAGTTACTAAC	4285
QY	241	ACACACCCCATTAACCAATTCTATCATGTGACATGTCTCAGCCGACTCTGAGAGTCTGTCAGC	300
Db	4286	ACACACCCCATTAACCAATAATCATGTGACATGTCTCAGCTGACCTGTGAGAGTCTGTCAGC	4345
QY	301	AGCACCTTGGGTGCTGTGTGGGTGGGTGCTTGGACACTCTGGCTGTGGTATTTGCTTGACAA	360
Db	4346	AGCACCTTGGGTGCTGTGTAGGGGAGGCTCTTGACACTCTTGGCCGGTATTTGCTTGACAA	4405
QY	361	GGCAGCGTGTGCATTGTGGGTGAGATCATCTTGTCCGGGCGGCGCGCTATTGTGTTCCGAC	420
Db	4406	GGCAGCGTGTGCATTGTGGGCGAGATCATCTTGTCCGGAAAGCGCGCATTCATCCCGAC	4465
QY	421	AGGGAAGTCTCTACACGAGATTGATGAGATGGAAGATGTCCGGTGTGCACCTTCCTTAC	480
Db	4466	AGGGAAGTCTTTCATCCGGGAGTTGTGATGAGATGGAAGATGTCCGCTCACACTTCCTTAC	4525
QY	481	ATCGAGCAGGGAATGCACTTGTCCGAGCAGTTCACGAAAAAGCGTCTCGGGTTGCTGAC	540
Db	4526	ATCGAACCAGGGAATGCACTGTCCGAAACAATTCAAACAGAAAGCAATTCGGGTGTCTGCAA	4585
QY	541	ACAGCCACCAAGCAAGCGAGGCGCTGTCCCTGTGTGAGTGTCCAATGGCGACCTT	600
Db	4586	ACAAGCCACCAAGCAAGCGAGGCTGTCTCTCCGTGTGAAATTCAAATGGCGGACCTTC	4645

Db 4706 TTGTCACACTGCTGGCGAACCCCGGATAGACATCATGATGGCATTCACAGCCTCTATC 4765  
 Qy 721 ACTAGCCCGCTACCAOCCCAATCTAACCTCTGCTTAAATCCTGGGAGGATGGGTAGCC 780  
 Db 4766 ACCAGCCCGCTCACACACCCAAATACCTCTGTTTAAATCTCGGGGGGATGGGTAGCC 4825  
 Qy 781 GCCCAACTGCTGCCCGCAGTGGCTTTCAGCTTTCGTAGGAGCGCGGCAATGTGCTGTGCG 840  
 Db 4826 GCCCACTGTGCTCTCCAGGCGTCTTCTCTTCTTGAAGCGCGGCAATGCTGAGAGCG 4885  
 Qy 841 GCTGTGGCAGCATAGGCTTTGGGAGGTGCTTGGACATCTTGGCGGCTATGAGACA 900  
 Db 4886 GCTGTGGCAGCATAGGCTTTGGGAGGTGCTTGGACATTTTGGCAGGTTATGAGACA 4945  
 Qy 901 GGAGTGGCAGGCGGCGCTCTGTGAGCTTTAAGGTATGAGCGGCGAATGSCCTCACCGAG 960  
 Db 4946 GGGGTGGAGGCGCGCTCTGTGCTTTAAGGTATGAGCGGCGAATGSCCTCACCGAG 5005  
 Qy 961 GACCTGTTAATTACTCTCCCTGTCATCTCTCTCTGCTGCTGCTGCTGTGCTGGGGTCTGT 1020  
 Db 5006 GACCTGTTAATTACTCTCCCTGTCATCTCTCTCCCTGCGCGCTTAATGCTGTGGGGTCTGT 5065

Db 5066 TGGCAGCATACTGCGTCGGCACGTGGGCCCGGGGGGGGGGCTGTCCAGTGGATGAC 5125

5126 CGGCTGATAGCGTTCCTTCGGGGGTACCAAGTCTTCCCGACGACTATGGCTGAG 5185  
1141 AGCAGCGCCGAGCAGCTGTCACTCAGATCCCTCCGACCTTACTATCACCACACTCTTG 1200  
5186 AGCAGCGTGCAGACGCTGACTCAAGATCCTCTTGTGCTTTACCACTACCTCACTGCTG 5245  
1201 AAGAGCTCCACCACTGGATTAAAGAGGAACTGTCTCCACGCGCTCTCGGCTCGTGCTA 1260  
5246 AAGAGGCTTACCACTGGATCAACAGAGAACTGCTCTCCACGCGCATCTCGGCTCGTGCTA 5305



QY 1261 AGGAGTGTGGGACGTGATATGACAGATTGGCTGACTTCAAGACCTGCTCAGTCC 1320  
DB 5306 AGAGATGTTGGGATGGATATGATACAGGTGTGATGATTTCAAGACCTGGCTCAGTCC 5365  
QY 1321 AAGCTCTGCGGGATTACCGGGAGTCCCTTTTCTCATGCGCAAGTGGGTCAAGGGG 1380  
DB 5366 AAGCTCTGCGGGATTGCGGGAGTCCCTTCTCATGTCAGTGGGTCAAGGGG 5425  
QY 1381 GTTGGCGGGGAGACGGCATCATGACAGACCACTGTCATGTGAGAGACAGATCACCGG 1440  
DB 5426 GTTGGCGGGGAGACGGCATCATGACAAACCACTGTCATGTGAGAGACAGATCACCGG 5485  
QY 1441 CATGTCAAAAACGGTTCATGAGAGATGTTGGGCTTAAGAAGCTGTAGTAAATGATGGCAT 1500  
DB 5486 CATGTCAAAAACGGTTCATGAGAGATGTTGGGCTTAAGAAGCTGTAGTAAATGATGGCAT 5545  
QY 1501 GGAACATTTCCCATCAACGATACACAGGGGCGCTGACGCGCTTCCCGACGGCCAAAC 1560  
DB 5546 GGAACATTTCCCATCAACGATACACAGGGGCGCTTCCCGACGGCCAAAT 5605  
QY 1561 TATTCAGGGGCGGTGGGGGGGCTGAGAGAGTGAAGGTTACGGGGGTGGG 1620  
DB 5606 TATTCAGGGGCGGTGGGGGGGCTGAGAGAGTGAAGGTTACGGGGGTGGG 5665  
QY 1621 GATTTCACTAGCTGACGACATGACCACTGAACAAGTAAATGCCGTGCGAGTTCCA 1680  
DB 5666 GATTTCACTAGCTGACGAGCATGACCACTGAACAAGTAAATGCCGTGCGAGTTCCG 5725  
QY 1681 GCCCGGAAATTTCTTCAACAGAGTGAATGGGGTGGCTGACAGTACGCTCCGGCGTGC 1740  
DB 5726 GCCCGGAAATTTCTTCAACAGAGTGAATGGGGTGGCTGACAGTACGCTCCAGGCGTGC 5785  
QY 1741 AAAACCTCTCTACGGGAGAGGTGACATTCACAGTGGGGCTCAACCAATTCCTGGTTGG 1800  
DB 5786 AAAACCTCTCTACGGGAGAGGTGACATTCCTGGGGCTCAATCAATACCTGGTTGG 5845  
QY 1801 TCGAGACTCCATGCGAGCCGCAACCGGATGTAGCATGTGCTCACTTCAATGCTCACCGAC 1860  
DB 5846 TCGAGACTCCATGCGAGCCGCAACCGGAGTGAAGTGTCTCAATGCTCACCGAC 5905  
QY 1861 CCGTCCCAATCAACAGACAGACGAGCTGAAGCGAGCTGGCCAGGGGGTCTCCCTCTCC 1920  
DB 5906 CCGTCCCAATCAACAGACAGACGAGCTGAAGCGAGCTGGCCAGGGGGATCTCCCTCTCC 5965  
QY 1921 TTGGCAGAGCTTCACTAGCCAGTGTCTGCGCTTCTCCGAAAGCGACATCACTTACC 1980  
DB 5966 TTGGCAGAGCTTCACTAGCCAGTGTCTGCGCTTCTCCGAAAGCAATGCACTTACC 6025  
QY 1981 CAAATGACTTCCCAAGCTGACCTCATGAGGCGCAACCTCTGTGGCGGACATGAGATG 2040  
DB 6026 CCGTCACTTCCCGGAGCGCTGACCTCATGAGGCGCAACCTCTGTGGCGGAGATG 6085  
QY 2041 GCGGGGAGCATTACCGCGGTGAGTCAAGAAACAAAGTATGTAATCTTGAATCTTTGAC 2100  
DB 6086 GCGGGGAGCATTACCGCGGTGAGTCAAGAAATPAAGTATGTAATTTTGAATCTTTGAC 6145  
QY 2101 CCGCTCCGAGCGGAGAGATGAGCGGAAAGTCCCTCCCGGAGATCCGCGGAGG 6205  
DB 6146 CCGCTCCGAGCGGAGAGATGAGCGGAAAGTATCCGTTCCCGGAGATCCGCGGAGG 6265  
QY 2161 TCCAAAGAAATTCACACAGAGATGCGGAGTGGGACGCGCGGATTTACAAACCTTCCGCT 2220  
DB 6206 TCCAAAGAAATTCCTTGAAGAGTCCCATATGAGGACGCGCGGATTTACAAACCTTCCGCT 6265  
QY 2221 CTGAGTCTTGAAGGCGCGGAGCTACGTCCTTCAATGAGTACATGGGTGCCACTGCCA 2280  
DB 6266 TTAGATCTTGAAGGAGACCGGAGCTACGTCCTTCAATGAGTACACGGGTGCCACTGCCG 6325  
QY 2281 CCTACTAAGACCTCTTATACCACTCCACGAGAGAGAGACATGTTGTCTGACAGAA 2340  
DB 6326 CCGCTTAAGGCGCGCTTCAATACCACTCCACGAGAGAGAGACGTTGTCTGACAGAA 6385  
QY 2341 TCACCGGTGTCTTCTGCGGTGAGGCTTTCACAAAGGCTTTCGATGCTCCGAACG 2400

DB 6386 TCTACCGGTGTCTTCTGCGGTGAGGAGCTGCGCAACAACCTTCCGAGCTCCGATG 6445  
QY 2401 TGGGCGGTGAGACAGCGGCAACGGCAACGCGCTCCGCAACAACCTTCCGAGCGGAG 2460  
DB 6446 TGGGCGGTGAGACAGCGGCAACGGCAACGCGCTCCGCAACAACCTTCCGAGCGGAG 6505  
QY 2461 GCAGATCTGAGGTTGAGTGTATTCCTTCATGCCCCCTTGAAGGGAGAGCGGGGAG 2520  
DB 6506 GCGGATCCGAGCGTTGAGTGTATTCCTTCATGCCCCCTTGAAGGGAGAGCGGGGAG 6565  
QY 2521 CCGGATCTGAGGAGCGGGCTTGTGTCTACCGTGAAGTGAAGGCGGTTAGAGCTGTCT 2580  
DB 6566 CCGGATCTGAGGAGCGGGCTTGTGTCTACCGTGAAGTGAAGGCGGTTAGAGAGCTGTCT 6625  
QY 2581 TGCTGTGATGTCTTCAACATGAGACAGGCGCTGATACCGCAATGCGCTGGAGGAA 2640  
DB 6626 TGCTGTGATGTCTTCAACATGAGACAGGCGCTGATACCGCAATGCGCTGGAGGAA 6685  
QY 2641 AGCAAGCTCCCATCAACGCTTGAAGCACTTTTGTGCTGCTCAACCAATGTCTAC 2700  
DB 6686 ACCAAGCTCCCATCAATGATGATGAGCACTTTTGTGCTGCTCAACCAATGTCTAC 6745  
QY 2701 GCTACCAATCCCGACGCGCAACGAGCGGCAAGAGGTCACTTTGACAGACTGCA 2760  
DB 6746 GCTACCAATCCCGACGCGCAACGAGCGGCAAGAGGTCACTTTGACAGACTGCA 6805  
QY 2761 ATCTGGAAGATCACTACAGAGAGTGTCTCAAGAGATGAAGCGGAGCGTCCACATG 2820  
DB 6806 ATCTGGAAGATCACTACAGAGAGTGTCTCAAGAGATGAAGCGGAGCGTCCACATG 6865  
QY 2821 AAGGCTAAGCTTCTATCAGTGAAGAGAGCTGCAAGCTGACGCGCCCATTCGCGCAA 2880  
DB 6866 AAGGCTAAGCTTCTATCAGTGAAGAGAGCTGCAAGCTGACGCGCCCATTCGCGCAA 6925  
QY 2881 TCTAAATTTGGCTATGAGGCAAGAGAGTCCGGAACCTATCCAGCAAGGCTTAAACAC 2940  
DB 6926 TCTAAATTTGGCTATGAGGCAAGAGAGTCCGGAACCTATCCAGCAAGGCTTAAACAC 6985  
QY 2941 ATCCGCTCCGTGTGAGAGAGCTTGTGAGAGACCTGGAACCAATGACACACATC 3000  
DB 6986 ATCCGCTCCGTGTGAGAGAGCTTGTGAGAGACCTGGAACCAATGACACACATC 7045  
QY 3001 ATGGCAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGAGCGGCAAGCGAGCTCCG 3060  
DB 7046 ATGGCAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGAGCGGCAAGCGAGCTCCG 7105  
QY 3061 CTATGCTTCCAGACTTGGGGGTCCGTGTGTGAGAGAAATGAGCCCTTATGAGCGT 3120  
DB 7106 CTATGCTTCCAGACTTGGGGGTCCGTGTGTGAGAGAAATGAGCCCTTATGAGCGT 7165  
QY 3121 GTCTCAACCTCCCTCAAGCTGATGAGGCTCTGTAGAGATTCAGATTTCTCTGGA 3180  
DB 7166 GTCTCAACCTCCCTCAAGCTGATGAGGCTCTGTAGAGATTCAGATTTCTCTGGA 7225  
QY 3181 CAGCGGGTGAAGTCTGTGTAACGCTGGAATCAAGAAAGCCCTTATGAGGCTTGA 3240  
DB 7226 CAGCGGGTGAAGTCTGTGTAATGCTGGAAGCGAAGAAATGCTTATGAGGCTTGA 7285  
QY 3241 TATGACACCGGCTTTTGAATCAACATGATGAGATGATCGTGTAGAGAGTGA 3300  
DB 7286 TATGACACCGGCTTTTGAATCAACATGATGAGATGATCGTGTGTAGAGAGTGA 7345  
QY 3301 ATTATCAATGTTGATCTTGGCCCCCGAAGCGACAGGCGATTAAGTGTCTCACAG 3360  
DB 7346 ATTATCAATGTTGATCTTGGCCCCCGAAGCGACAGGCGATTAAGTGTCTCACAG 7405  
QY 3361 CCGCTTATATGAGGGGTCCCTGACTAATTAAGGCGAGAACTGAGGCTATGCGCG 3420  
DB 7406 CCGCTTATATGAGGGGTCCCTGACTAATTAAGGCGAGAACTGAGGCTATGCGCG 7465  
QY 3421 TGCAGCGAGGCGCTGTGACGATGAGTCCGATTAATACCTCAATGTTACTTGAAG 3480

Db 7466 TGCCTGCGAGCGGTGTAAGTGAACGACACCTGCGGTAAATACCTCATGTATCTTGAAG 7525  
QY 3481 GCCTTGACAGCCTGTGAGCTGCAAGCTCCAGACCTGACATGTCTGTGTCGGAGAC 3540  
Db 7526 GCGCTGCGGCTGTGTGAGTGTGGAAGCTTCAGACCTGACATGTCTGTATGCGAGAC 7585  
QY 3541 GGCCTGTGCTTATCTGTGAGAGCGGGAACCCAGAGAGCGGCGAGCTTACGAGTC 3600  
Db 7586 GACCTGTGCTTATCTGTGAAAGCGGGAACCCAGAGAGCGAGGCGAGCTTACGAGTC 7645  
QY 3601 TTACAGAGAGCTTATGCTAGTACTCTGCCCCCCCCCGGAGACCCGCGCAACCAAGATAC 3660  
Db 7646 TTACAGAGAGCTTATGCTAGTACTCTGCCCCCCCCCGGAGAGCCGCGCAACCAAGATAC 7705  
QY 3661 GACCTGAGTGTATATCATCATGCTCCTCCATGTGTGCGTCCGACGACATGATCTGAC 3720  
Db 7706 GACTTGAAGTTGATACATATCTCTCTCAATGTGTCAGTCCGACGACATGATCTGAC 7765  
QY 3721 AAAAGGTATATCTACTCAACCCGTAACC 3749  
Db 7766 AAAAGGTATATCTACTCAACCCGTAACC 7794

RESULT 11  
US-10-029-907-7  
; Sequence 7, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802)...(8407)  
US-10-029-907-7

Query Match 87.7%; Score 3288.2; DB 4; Length 8638;  
Best Local Similarity 92.3%; Pred. No. 0;  
Matches 3461; Conservative 0; Mismatches 288; Indels 0; Gaps 0;  
QY 1 TGGAGAGGCGCTTTCAAGGCTCAACCAAGTGAATCCCACTTCTGTCCTCAAAACAAAG 60  
Db 4046 TGGAGAGGCGCTTTCAAGGCTCAACCAAGTGAATCCCACTTCTGTCCTCAAAACAAAG 4105  
QY 61 CAGGACAGAGACACTTCCCTACTGATGAGCGTACAGAGTACTGATGAGCGTAAAGGAGCC 120  
Db 4106 CAGGACAGAGACACTTCCCTACTGATGAGCGTACAGAGTACTGATGAGCGTAAAGGAGCC 4165  
QY 121 CAGGACCCACCTCATCATGAGATCAATGTGAAGTGTCTCATAGGCTTAAAGCTTACT 180  
Db 4166 CAGGCTCACTCATCATGAGATCAATGTGAAGTGTCTCATAGGCTTAAAGCTTACT 4225  
QY 181 CTGCGCGGCGCAACACCTTGTGTATAGCTGGAGCGCTCAAAACGAGGTCAACCTTC 240  
Db 4226 CTGCGCGGCGCAACACCTTGTGTATAGCTGGAGCGCTTCAAAACGAGGTCAACCTTC 4285  
QY 241 ACACACCCCAATTAACCAATTCATATGAGCATGATGACGACGCTGAGGTCTGACG 300  
Db 4286 ACACACCCCAATTAACCAATTCATATGAGCATGATGACGATGCTGAGGTCTGACG 4345  
QY 301 AGACATGAGGTGTGCTGAGCGGCGCTTTCAGACTCTGAGTGTGCTTGTGACACA 360

Db 4346 AGCACCCTGAGTGTGATGAGCGGAGTCTTACAGCTCTGAGTGTGCTATTTGCTGACACA 4405  
QY 361 GGCACGCTGTATCTGTGAGTGAATCATCTTGTCCGAGCGCGCTATTTGTTCCGAC 420  
Db 4406 GGCACGCTGTATCTGTGAGTGAATCATCTTGTCCGAGCGCGCTATTTGTTCCGAC 4465  
QY 421 AGGGAAGTCTTACACAGAGTTCATGATGAGAAAGTGTGCGCTCCGACCTCCCTTAC 480  
Db 4466 AGGGAAGTCTTACACAGAGTTCATGATGAGAAAGTGTGCGCTCCGACCTCCCTTAC 4525  
QY 481 ATCGACAGGGGAATTCAGCTGCGCGAGCATTTCAAGCAAAAGCGCTGAGGTGTGAG 540  
Db 4526 ATCGACAGGGGAATTCAGCTGCGCGAGCATTTCAAGCAAAAGCGCTGAGGTGTGAG 4585  
QY 541 ACAGCACCAAGCAAGCGAGGCGCTGCTCCGATGTGAGTCCAAAGTGTGAGAGCTT 600  
Db 4586 ACAGCACCAAGCAAGCGAGGCTGCTCCGATGTGAGTCCAAAGTGTGAGAGCTT 4645  
QY 601 GAGACCTTGTGGGCGAACAATGTGAACTTATCAGCGGATACATGATCTTACAGAGC 660  
Db 4646 GAGACCTTGTGGGCGAAGCATGTGAACTTATCAGCGGATACATGATCTTACAGAGC 4705  
QY 661 TTGTCACTCTGCGCGGGAATCCGCGATGATCATGATGAGGCTTCAAGCTCTGTC 720  
Db 4706 TTGTCACTCTGCGCGGGAATCCGCGATGATCATGATGAGGCTTCAAGCTCTTATC 4765  
QY 721 ACTAGCGCGCTACACCAACCATCTACCTCTCTTAAATCTCTGAGGAGATGAGTAC 780  
Db 4766 ACTAGCGCGCTACACCAACCATCTACCTCTCTTAAATCTCTGAGGAGATGAGTAC 4825  
QY 781 GCCCACTGCTCCCCCAAGTCTGCTTCAAGTTCATGAGGCGCGCATTCCTGCTG 840  
Db 4826 GCCCACTGCTCCCCCAAGTCTGCTTCAAGTTCATGAGGCGCGCATTCCTGCTGAG 4885  
QY 841 GCTGTGAGAGATAGGCTTGGGAAAGTGTGAGCATCTTGGCGGCTATGAGAGA 900  
Db 4886 GCTGTGAGAGATAGGCTTGGGAAAGTGTGAGCATCTTGGCGGCTATGAGAGA 4945  
QY 901 GAGTGTGAGAGGCGCTGCTGAGCTTAAAGTTCATGAGGCGGAAATGCTTCAACGAG 960  
Db 4946 GAGTGTGAGAGGCGCTGCTGAGCTTAAAGTTCATGAGGCGGAAATGCTTCAACGAG 5005  
QY 961 GACCTGTAACTTACTCTCTGCAATCTCTCTCTGAGTGTGCTGAGGCTGCTG 1020  
Db 5006 GACCTGTAACTTACTCTCTGCAATCTCTCTCTGAGTGTGCTGAGGCTGCTGAG 5065  
QY 1021 TGGCAGCGATATCTGCTGAGGAGCGTGTCCAGAGGAGGAGGCTGTGAGTGAATGAC 1080  
Db 5066 TGGCAGCGATATCTGCTGAGGAGCGTGTCCAGAGGAGGAGGCTGTGAGTGAATGAC 5125  
QY 1081 CGGCTGATGAGCTTCCCTCGCGGAGTAACTGTTTCCCAAGCATATGTGCCAGAG 1140  
Db 5126 CGGCTGATGAGCTTCCCTCGCGGAGTAACTGTTTCCCAAGCATATGTGCCAGAG 5185  
QY 1141 AGGACGCGAGACAGTGTCACTGATCTCTGCACTTATCTATCAACCACTGTTG 1200  
Db 5186 AGGACGCGAGACAGTGTCACTGATCTCTGCACTTATCTATCAACCACTGTTGAG 5245  
QY 1201 AAGAGCTCAACAGATGATTAACAGAGCTGCTCAAGGCTGCTGAGCTGTGAGTGA 1260  
Db 5246 AAGAGCTCAACAGATGATTAACAGAGCTGCTCAAGGCTGCTGAGCTGTGAGTGA 5305  
QY 1261 AAGGATGTTTGGGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
Db 5306 AAGGATGTTTGGGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5365  
QY 5366 AAGCTCTCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5425  
Db 1381 GTCTGCGGAGAGACGCGCATATGACAGCACTGCTATGTGAGACATGATCACCGGA 1440  
QY 5426 GTCTGCGGAGAGACGCGCATATGACAGCACTGCTATGTGAGACATGATCACCGGA 5485

QY 1441 CATGTCAAAAACGGTTCATGAGATCGTTGGCCTTAAGAAGCTGTAGTAACATGTGGCAT 1500  
DB 5486 CATGTAAAAACGTTTCATGAGGATCGTGGGCTTAAGACCTGTAGTAACATGTGGCAT 5545  
QY 1501 GGAACATTTCCCATCAACGCAATACCAACGGGCTCTGACAGCCTTCCCAAGCCCAAC 1560  
DB 5546 GGAACATTTCCCATCAACGCAATACCAACGGGCTCTGACAGCCTTCCCAAGCCCAAC 1560  
QY 1561 TATTCAGGAGCGGTGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1620  
DB 5606 TATTCAGGAGCGGTGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1620  
QY 1621 GATTTCCACTACGTGACGAGCATGACCACTGAACAAGTAATGCGGTCCAGATTCA 1680  
DB 5666 GATTTCCACTACGTGACGAGCATGACCACTGAACAAGTAATGCGGTCCAGATTCA 1680  
QY 1681 GCCCCGGAATTTCTCAACGAATGATGGGGTGGGCTGACAGGTAACGCTCCGCGTGC 1740  
DB 5726 GCCCCGGAATTTCTCAACGAATGATGGGGTGGGCTGACAGGTAACGCTCCGCGTGC 1740  
QY 1741 AAAACCTCTCTACGGGAGGAGGATCATATTCAGGTGGGTGTAACCAATATCTGTGGG 1800  
DB 5786 AAAACCTCTCTACGGGAGGAGGATCATATTCAGGTGGGTGTAACCAATATCTGTGGG 1800  
QY 1801 TCGAGCTCCCATGACGAGCCCGAACCAGATGTAGCAGTGCTCACTTCCATGCTCACCGAC 1860  
DB 5846 TCGAGCTCCCATGACGAGCCCGAACCAGATGTAGCAGTGCTCACTTCCATGCTCACCGAC 1860  
QY 1861 CCTCCACATCAACAGCAGAGACGAGTAAAGCAGGCTGGCCAGGGGCTCTCCCTCC 1920  
DB 5906 CCTCCACATCAACAGCAGAGACGAGTAAAGCAGGCTGGCCAGGGGCTCTCCCTCC 1920  
QY 1921 TTGGCAGACTTTCACTAGCTAGCCAGTTGTCTGGCCTTCTCGAAGGCAATCATTTAC 1980  
DB 5966 TTGGCAGACTTTCACTAGCTAGCCAGTTGTGTGCGCTTCTCGAAGGCAATCATTTAC 1980  
QY 1981 CAAATGACTTCCACAGCGTGAACCTCATGAGGCAACCTCTGAGGGGCAATGAGATG 2040  
DB 6026 CCAATGACTTCCACAGCGTGAACCTCATGAGGCAACCTCTGAGGGGCAATGAGATG 2040  
QY 2041 GCGGGGAGCATTAACCGCGTGAAGTACAGAAACAGTGTATCTCTGACCTTTTCAC 2100  
DB 6086 GCGGGGAGCATTAACCGCGTGAAGTACAGAAACAGTGTATCTCTGACCTTTTCAC 2100  
QY 2101 CCGCTCCAGCGAGGAGATGAGCGGAAAGTGTCTCCGCGGAGATCTCTCGGAAA 2160  
DB 6146 CCGCTCCAGCGAGGAGATGAGCGGAAAGTGTCTCCGCGGAGATCTCTCGGAAA 2160  
QY 2161 TCCAGAAATTTCCACACAGCATGCGCGCATGAGGACGCGCGGATTAACAACCTCCGCTG 2220  
DB 6206 TCCAGAAATTTCCACACAGCATGCGCGCATGAGGACGCGCGGATTAACAACCTCCGCTG 2220  
QY 2221 CTGAGTCTGGAAGGCGCGGACTAGCTCCCTCAAGTGTATCATGGTGGCCACTGCA 2280  
DB 6266 TTAGAGTCTGGAAGGCGCGGACTAGCTCCCTCAAGTGTATCATGGTGGCCACTGCA 2280  
QY 2281 CCTACTAGACCCCTCTATACCACTCCACAGGAGAGAGACAGTTGTCTGACAGAA 2340  
DB 6326 CCTACTAGACCCCTCTATACCACTCCACAGGAGAGAGACAGTTGTCTGACAGAA 2340  
QY 2341 TCCACCGTGTCTTTGCTGCGGAGGCTTCCCAAAAGCTTTGGTGTAGTCTGCAACG 2400  
DB 6386 TCTACCGTGTCTTTGCTGCGGAGGCTTCCCAAAAGCTTTGGTGTAGTCTGCAACG 2400  
QY 2401 TCGGCGGTGACAGCGGACAGGCAACGCGCTCTGACCAACCTCCGACGACGCGGAG 2460  
DB 6446 TCGGCGGTGACAGCGGACAGGCAACGCGCTCTGACCAACCTCCGACGACGCGGAG 2460  
QY 2461 GAGAGATCTGAGCTGTAGTCTATCTCTCATGCCCCCTTGAAGGAGAGCGCGGAGAC 2520  
DB 6506 GCGGAGATCTGAGCTGTAGTCTATCTCTCATGCCCCCTTGAAGGAGAGCGCGGAGAT 2520

QY 2521 CCGATCTCAGGACGCGGTCTTGCTTACCGTGAAGTGAAGGCGCGGTGAGACGTCTGC 2580  
DB 6566 CCGATCTCAGGAGCGGCTTGCTTACCGTGAAGTGAAGGCGCGGTGAGACGTCTGC 2580  
QY 2581 TGTGCTGATGTCTTACCATGAGACGCGCTCTGATCAACGCTATGCGTGGAGGAA 2640  
DB 6626 TGTGCTGATGTCTTACCATGAGACGCGCTCTGATCAACGCTATGCGTGGAGGAA 2640  
QY 2641 AGCAAGTCCCATCAACGCGTGTGACCACTCTTGTGCGGTGACCAACATGCTCAT 2700  
DB 6686 AGCAAGTCCCATCAACGCGTGTGACCACTCTTGTGCGGTGACCAACATGCTCAT 2700  
QY 2701 GCTACCAATCCCGACGCGCAAGCCAGCGGAGAGAGTCACTTTGACAGACTGCA 2760  
DB 6746 GCTACCAATCCCGACGCGCAAGCCAGCGGAGAGAGTCACTTTGACAGACTGCA 2760  
QY 2761 ATCTGAGCATCACTACCAAGACGTGCTCAAGAGATGAAGCGGAGCGTCCACGTT 2820  
DB 6806 ATCTGAGCATCACTACCAAGACGTGCTCAAGAGATGAAGCGGAGCGTCCACGTT 2820  
QY 2821 AAGGCTAAGCTTCTATCAGTGAAGAGCGTCAAGCTGACGCGCCCAATTCGCGCAA 2880  
DB 6866 AAGGCTAAGCTTCTATCAGTGAAGAGCGTCAAGCTGACGCGCCCAATTCGCGCAA 2880  
QY 2881 TCTAAATTTGGCTATGAGGCAAGACGTCGGAACCTTACGACAGGCAATTAACAC 2940  
DB 6926 TCTAAATTTGGCTATGAGGCAAGACGTCGGAACCTTACGACAGGCAATTAACAC 2940  
QY 2941 ATCCGCTCGTGTGAGAGACTTGTGGAAGACCTGAACCAATTTGACACCAACATC 3000  
DB 6986 ATCCGCTCGTGTGAGAGACTTGTGGAAGACCTGAACCAATTTGACACCAACATC 3000  
QY 3001 ATGCAAAAATGAGGTTTTCTGCGTCCACCAAGAGAGAGGCGCGGACGCTCGC 3060  
DB 7046 ATGCAAAAATGAGGTTTTCTGCGTCCACCAAGAGAGAGGCGCGGACGCTCGC 3060  
QY 3061 CTATGATTTCCAGACTTGGGAGGCTCGTGTGAGAGAAATGAGGCTCTATGACGTG 3120  
DB 7106 CTATGATTTCCAGACTTGGGAGGCTCGTGTGAGAGAAATGAGGCTCTATGACGTG 3120  
QY 3121 GTCTCACCCCTCTCAGGCTGTGATGAGCTCTCTGATCGATTTCTCTGGA 3180  
DB 7166 GTCTCACCCCTCTCAGGCTGTGATGAGCTCTCTGATCGATTTCTCTGGA 3180  
QY 3181 CAGCGGCTGAGTTCTGCTGTAACGCTGGAATCAAGAGACCCCTATGAGGCTTTGCA 3240  
DB 7226 CAGCGGCTGAGTTCTGCTGTAACGCTGGAATCAAGAGACCCCTATGAGGCTTTGCA 3240  
QY 3241 TATGACACCGGCTTTGATCAACAGTCACTGAGATGATGATGATGATGATGATGATG 3300  
DB 7286 TATGACACCGGCTTTGATCAACAGTCACTGAGATGATGATGATGATGATGATGATG 3300  
QY 3301 ATTATCAATGTTGATCTTGGCCCCCGAGGACAGACAGGCTATAGCTGACAGAG 3360  
DB 7346 ATTATCAATGTTGATCTTGGCCCCCGAGGACAGACAGGCTATAGCTGACAGAG 3360  
QY 3361 CGGCTTATATGAGGAGTCCCTGATCTAATTAAGAGGACAGCTGAGTATGCGCG 3420  
DB 7406 CGGCTTATATGAGGAGTCCCTGATCTAATTAAGAGGACAGCTGAGTATGCGCG 3420  
QY 3421 TGCAGGAGAGGAGGAGTGTGACGATGAGTGGGTAATACCTCAATGTTACTTGAAG 3480  
DB 7466 TGCAGGAGAGGAGGAGTGTGACGATGAGTGGGTAATACCTCAATGTTACTTGAAG 3480  
QY 3481 GCTCTGACGCTGTGAGCTGCAAGCTCCAGACTGACAGATGCTGTGTGCGGAGAC 3540  
DB 7526 GCTCTGACGCTGTGAGCTGCAAGCTCCAGACTGACAGATGCTGTGTGCGGAGAC 3540  
QY 3541 GGCCTTGTGTTATGTGAGAGCGCGGAAACCAAGAGAGCGCGGACGCTTACGAGTC 3600  
DB 7586 GACCTTGTGTTATGTGAGAGCGCGGAAACCAAGAGAGCGCGGACGCTTACGAGTC 3600  
QY 3601 TTACAGGAGGCTATGATGATGATCTGCCCCCCCCCGGAGACCGGCCCCCAACCAATAC 3660

|||||  
Db 7646 TTCACGAGGCTATGACTAGATCTGCCCCCTGGGAGCCGCCCAACCAAGATAC 7705  
QY 3661 GACCTGGAATTGATATCATCATCTCTCCAAATGTTCGGTCGGGACAGATGATTTGGC 3720  
Db 7706 GACTTGGAATGTGATTAATCATCATGCTCTCCAAATGTTCAGTCGGGACAGATGATCTGGC 7765  
QY 3721 AAAAGGTATTAATCACTCAACCCGTGACCC 3749  
Db 7766 AAAAGGTATTAATCACTCAACCCGTGACCC 7794  
  
RESULT 12  
US-10-029-907-25  
; Sequence 25, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PaetSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802) ... (8407)  
US-10-029-907-25  
  
Query Match 87.7%; Score 3288.2; DB 4; Length 8638;  
Best Local Similarity 92.3%; Pred. No. 0;  
Matches 3461; Conservative 0; Mismatches 288; Indels 0; Gaps 0;  
  
QY 1 TGGGAGGGCGCTTTCACAGGCTCACCCAGTGAATGCCACTTCTGTCACAAAG 60  
Db 4046 TGGGAGGGCGCTTTCACAGGCTCACCCAGTGAATGCCACTTCTGTCACAAAG 4105  
QY 61 CAGGACGAGACAACTTCCCTACCTGTGGCGGTACCAAGCTACTGTGTGGCTAGAGGC 120  
Db 4106 CAGGACGAGACAACTTCCCTACCTGTGTGGCGGTACCAAGCTACTGTGTGGCGGTAG 4165  
QY 121 CAGGACCCACCTTCATCATGTGAATGGAAGTGTCTCATACGGCTTAAGCCTACT 180  
Db 4166 CAGGCTCCACTTCATCATGTGGACCAATGTGAAGTGTCTCATACGGCTTAAGCCTACT 4225  
QY 181 CTGCGCGGGCCAAACCCCTTGTGTATAGGCTGGAGCCGCTCCAAAACGAGTCAACCTTC 240  
Db 4226 CTGCGCGGGCCAAACCCCTTGTGTATAGGCTGGAGCCGCTCCAAAACGAGTCAACCTTC 4285  
QY 241 ACACACCCCTTAACCAATTCATCATGTGACATGTCAAGCCGACTGAGGCTGTACAG 300  
Db 4286 ACACACCCCTTAACCAATTCATCATGTGACATGTGACATGTGAGCTGAGGCTGTACAG 4345  
QY 301 AGACCTGGGTGTGGTGGGCGGGGTCTTCGAGCTGTGGCTGATTCCTGAACA 360  
Db 4346 AGACCTGGGTGTGGTGGGCGGGGTCTTCGAGCTGTGGCTGATTCCTGAACA 4405  
QY 361 GGCAGCGTGTCAATTGTGGTAGAGATCTTGTCCGGGCGCGCTATTGTTCGAC 420  
Db 4406 GGCAGCGTGTCAATTGTGGTAGAGATCTTGTCCGGGAGCGCGCATTCATTCGAC 4465  
QY 421 AGGGAAGTCTCTTACCAAGAGTTCATGATGAGATGGAAGTGCAGCTTCCTTAC 480  
Db 4466 AGGGAAGTCTCTTACCAAGAGTTCATGATGAGATGGAAGTGCAGCTTCCTTAC 4525  
QY 481 ATGAGCAGGGAATGCACTCGCGAGCAATTCAAGAAAAAGCGCTCGGGTTGCTGAC 540

|||||  
Db 4526 ATCGAACAGGGAATCAGCTCCGCCAACATTCAAACGAAGCAATCGGGTGTGCA 4585  
QY 541 AAGGCAACCAACAGCGGAGCGCTGCTCCGATGTGAGATCCAAAGTGGAGCCCTT 600  
Db 4586 AAGGCAACCAACAGCGGAGCGCTGCTCCGATGTGAGATCCAAAGTGGAGCCCTT 4645  
QY 601 GAGACCTTCTGGGCGAACAACATGTGAATTCATCAAGCGGATACATTAAGTACAGGC 660  
Db 4646 GAGACCTTCTGGGCGAAGATATGTGAATTCATCAAGCGGATACATTAAGTACAGGC 4705  
QY 661 TTGTCCACTCTGCTGGGAATCCCGATGTGATCATGTATGCGTTCAACGCTCTGTC 720  
Db 4706 TTGTCCACTCTGCTGGGAATCCCGATGTGATCATGTATGCGTTCAACGCTCTGTC 4765  
QY 721 ACTAGCGGCTCACCAACCAATCTACCCCTCGCTTAACATCTGAGGAGATGGGTAGCC 780  
Db 4766 ACTAGCGGCTCACCAACCAATCTACCCCTCGCTTAACATCTGAGGAGATGGGTAGCC 4825  
QY 781 GCCCACTGCTCCCCAGTGTGCTTCAGCTTCTGATAGGCGCGCATTTGCTGTGCG 840  
Db 4826 GCCCACTGCTCCCCAGTGTGCTTCAGCTTCTGATAGGCGCGCATTTGCTGTGCG 4885  
QY 841 GCTGTGGCAGATAGGCTTGGGAAGTGTGTGACATTTGCGGCGCTATGAGCA 900  
Db 4886 GCTGTGGCAGATAGGCTTGGGAAGTGTGTGATATTGTCAGATATGAGCA 4945  
QY 901 GGAATGGCAGGCGGCTCGTGGCTTTAAGTCATGAGCGGGAATGCCCTCCACCGAG 960  
Db 4946 GGAATGGCAGGCGGCTCGTGGCTTTAAGTCATGAGCGGGAATGCCCTCCACCGAG 5005  
QY 961 GACTGTGTTAATCTTACTCCCTGCAATCTCTCTCTGAGCCCTGATGCTCGGAGTCG 1020  
Db 5006 GACTGTGTTAATCTTACTCCCTGCAATCTCTCTCTGAGCCCTGATGCTCGGAGTCG 5065  
QY 1021 TGGCAGCGATATCTGCTGGGACGCTGGGTCCAGGGAGGGGCTGTGAGTGAATAC 1080  
Db 5066 TGGCAGCGATATCTGCTGGGACGCTGGGTCCAGGGAGGGGCTGTGAGTGAATAC 5125  
QY 1081 CGGCTGATAGCTTGGCTCGGGGGGTAAACATGTTTCCCGACGACATATGTCAGAG 1140  
Db 5126 CGGCTGATAGCTTGGCTCGGGGGGTAAACATGTTTCCCGACGACATATGTCAGAG 5185  
QY 1141 AGGACGCGGACGAGTGTCACTAGATCTCTCCGACCTTACTATCAACCACTGTG 1200  
Db 5186 AGGACGCTGACGACAGTGTCACTAGATCTCTCTGATCTTCAATCACTAGCTGTG 5245  
QY 1201 AAGAGCTTCACAGTGAATTAACAGAGCTGTCCAGCCCTGCTCGGCTGTGCTTA 1260  
Db 5246 AAGAGCTTCACAGTGAATTAACAGAGCTGTCCAGCCCTGCTCGGCTGTGCTTA 5305  
QY 1261 AAGGATGTTGGGATCTGATATGACAGTGTGGCTGACTTCAAGACCTGCTCCAGTTC 1320  
Db 5306 AAGGATGTTGGGATCTGATATGACAGTGTGGCTGACTTCAAGACCTGCTCCAGTTC 5365  
QY 1321 AAGCTCTGCGCGAATTAACGGGAGTCCCTTTTCTCATGCGCAAGTGGATACAGGGG 1380  
Db 5366 AAGCTCTGCGCGAATTAACGGGAGTCCCTTTTCTCATGCGCAAGTGGATACAGGGG 5425  
QY 1381 GTCTGGCGGGAGACGGCATATGACCAACCTGTCTATGTGAGACACAGATCAACGGA 1440  
Db 5426 GTCTGGCGGGAGACGGCATATGACCAACCTGTCTATGTGAGACACAGATCAACGGA 5485  
QY 1441 CATGTCAAAAACGCTTCATGAGATGTTGGGCTTAAGACTGTATGAATGTGGCAT 1500  
Db 5486 CATGTCAAAAACGCTTCATGAGATGTTGGGCTTAAGACTGTATGAATGTGGCAT 5545  
QY 1501 GGAACATTCCTCATCAACGATACCAAGGGCCCTTGACGCGCTCCAGCGCCAAAC 1560  
Db 5546 GGAACATTCCTCATCAACGATACCAAGGGCCCTTGACGCGCTCCAGCGCCAAAC 5605  
QY 1561 TATTCAGGGCGCTGTGGGCTGTGAGAGTATGTTGAGAGTTACGCGGCTGGG 1620

D	5606	TATTTCTAGGGCGCCTTGCGCGGGTGGCTCTCTAGAGATGATCGTAGAGTTAACCGCACTGGGG	5665
Q	1621	GATTTTCCAATACTGATGACGACATGACCTGACCACTGACCAAGTAAATGCCCCGTGCAGAGTTTCCA	1680
D	5666	GATTTTCCAATACTGATGACGACGAGATGACCACTGACCAAGTAAATGCCCCGTGCAGAGTTTCCA	5725
Q	1681	GGCCCCGGAATTTCTTACAGAAATGGAATGGGGTGGCGGCTGACAGGTAAGCTTCCGGCGTGC	1740
D	5726	GGCCCCGGAATTTCTTACAGAAATGGAATGGGGTGGCGGCTGACAGGTAAGCTTCCGGCGTGC	5785
Q	1741	AAACCTCTCTACACGGGAGAGGATGACATTTCCAGAGTCCGGGCTCAACCAATATCTGGTTGGG	1800
D	5786	AAACCTCTCTCTACACGGGAGAGGATGACATTTCTGAGTCCGGGCTCAACCAATATCTGGTTGGG	5845
Q	1801	TTGCAAGCTTCCATGCGAGCCCGAACCGGATGATGACAGTGTCTCACTTCCATGCTCAACCGAC	1860
D	5846	TTGCAAGCTTCCATGCGAGCCCGAACCGGATGATGACAGTGTCTCACTTCCATGCTCAACCGAC	5905
Q	1861	CCCTTCCCACTTACACAGCAGACACGGCTTAAGCGCAGGCTGGCCAGGGGGGTCTTCCCCCTCC	1920
D	5906	CCCTTCCCACTTACACAGCAGACACGGCTTAAGCGCTAAGCGTGAAGCTGGCCAGGGGATCTTCCCCCTCC	5965
Q	1921	TTGGCAGAGCTCTTACGCTTACAGCAGTGTCTGCGGCTTCTCGAAGGGCGACATTAACATTACC	1980
D	5966	TTGGCAGAGCTCTTACGCTTACAGCAGTGTCTGCGGCTTCTCGAAGGGCGACATTAACATTACC	6025
Q	1981	CAAAATGACTTCCACAGACGCTGACCTCTATCGAGGCGCACTCTGTGGCGGACATGATGATG	2040
D	6026	CGTATGACTTCCACAGACGCTGACCTCTATCGAGGCGCACTCTGTGTGGCGGACATGATGATG	6085
Q	2041	GGCGGGGACATTTACCCCGGTGAGTCTAGAGAACAAAGTATCTTGGACTCTTTTGAC	2100
D	6086	GGCGGGGACATTTACCCCGGTGAGTCTAGAGAACAAAGTATCTTGGACTCTTTTGAC	6145
Q	2101	CCGCTCCGAGGGGAGAGAGATGAGGGGGGAAAGTGTCCGTCCGGGGGAGAAATCTCTCGGAAA	2160
D	6146	CCGCTCCGAGGGGAGAGAGATGAGGGGGGAAAGTATCTTCCGGGGGAGAAATCTCTCGGAAA	6205
Q	2161	TTCCAAGAAATTTCCACACAGCAGTGGCCGACATGAGGACCGCCGATTAACACCTTCCGCTG	2220
D	6206	TTCCAAGAAATTTCCACACAGCAGTGGCCGACATGAGGACCGCCGATTAACACCTTCCGCTG	6265
Q	2221	CTGAGTCTTGAGAGCCCCCGGACTAGCTCCCTTCCAGTGTGATGATGGTGTCCCACTGCA	2280
D	6266	CTGAGTCTTGAGAGCCCCCGGACTAGCTCCCTTCCAGTGTGATGATGGTGTCCCACTGCA	6325
Q	2281	CCCTACTAGACCCCTCTCTATACACCTTCCACAGGAGGAAAGAGACAGTTGTTCTGACAGAA	2340
D	6326	CCCTACTAGACCCCTCTCTATACACCTTCCACAGGAGGAAAGAGACAGTTGTTCTGACAGAA	6385
Q	2341	TTCAACCGATCTTCTTGCGCCGTGCGGAGGCTTGACCAAAAGGCTTTTCGGTAGTCCGAAACCG	2400
D	6386	TTCAACCGATCTTCTTGCGCCGTGCGGAGGCTTGACCAAAAGGCTTTTCGGTAGTCCGAAACCG	6445
Q	2401	TCGGCCGTGACAGCGGACCGGCAACCGCCCCCTCTGACCAACCTTCCGACGACGCGGGA	2460
D	6446	TCGGCCGTGACAGCGGACCGGCAACCGCCCCCTCTGACCAACCTTCCGACGACGCGGGA	6505
Q	2461	GCAAGATCTGACGTTGAGTCTGATTTCTTCTCAATGCCCCCTTGAAGGGGAGACCGGGGAC	2520
D	6506	GCAAGATCTGACGTTGAGTCTGATTTCTTCTCAATGCCCCCTTGAAGGGGAGACCGGGGAT	6565
Q	2521	CCCGATCTCAGGACCGGCTTTGGCTCTACCGGATGAGGAGGCGCGGATGAGAGACGTCGTC	2580
D	6566	CCCGATCTCAGGACCGGCTTTGGCTCTACCGGATGAGGAGGCGCGGATGAGAGACGTCGTC	6625
Q	2581	TGCTGCTGATGTCTTACATGACAGGACGCTCTGATCAGCGCATGCGCTGCGGAGAAA	2640
D	6626	TGCTGCTGATGTCTTACATGACAGGACGCTCTGATCAGCGCATGCGCTGCGGAGAAA	6685
Q	2641	AGCAAGCTGCCCTATACAGCGGTGAGCACTTTTGTGCGGTCAACACATAGTGTCTAC	2700
D	6686	AGCAAGCTGCCCTATACAGCGGTGAGCACTTTTGTGCGGTCAACACATAGTGTGTCTAT	6745

QY	2701	ACTACCAATCCCGAGGCAAGCAGGCGGAGAAAGGACACTTTGACAGCTGAA	2760
Db	6746	GCTACACATCTCGAGCGCAAGCTCGGCGAAGAAAGTCACTTTGACAGCTGAG	6805
QY	2761	ATCTGACGATCACTACCAAGACGTGCTCAAGAGATGAAAGCGAAGCGCTCCAGTT	2820
Db	6806	GTCCTGACGACCACTACCGGAGCGTGCTCAAGAGATGAAAGCGAAGCGCTCCAGTT	6865
QY	2821	AAGGCTAAGCTTCTATCAGTAGAGAAAGCTTGCAAGCTGACGCCCCCAATTGGGCCAA	2880
Db	6866	AAGGCTAAACTCTTATCCGTGAGAGAAAGCCGTAAAGCTGACGCCCCCAATTGGGCCAA	6925
QY	2881	TCTAAATTGGCTATGGGGCAAAAGACGTCCGGAACCTATCCAGCAAGGCCATTAAACAC	2940
Db	6926	TCTAAATTGGCTATGGGGCAAAAGACGTCCGGAACCTATCCAGCAAGGCCATTAAACAC	6985
QY	2941	ATCCGCTCCGTGTGAGAGACTTGTGGAAGACACTGAACCAATTGACCAACCAATC	3000
Db	6986	ATCCGCTCCGTGTGAGAGACTTGTGGAAGACACTGAACCAATTGACCAACCAATC	7045
QY	3001	ATGCGAAAAATGAGGTTTCTGGGTCAACCAAGAGAGAGGCGCAGCCAGCCGCTGCG	3060
Db	7046	ATGCGAAAAATGAGGTTTCTGGGTCAACCAAGAGAGAGGCGCGCCAGCCAGCCGCTGCG	7105
QY	3061	CTTATCGGTTCACAGACTTGGGGGTCGCTGTGTGCGAGAAATGAGCCCTCTATGACGTG	3120
Db	7106	CTTATCGGTATCCAGATTGGGGGTTCTGTGTGTGCGAGAAATGAGCCCTTATGAGTGTG	7165
QY	3121	GTCCTCACCCCTCCCTCAGGCTGTGATAGGGCTCCTCTGACGGAATTCCTGCTCGGA	3180
Db	7166	GTCCTCACCCCTCCCTCAGGCGGTATAGGGCTCTTCAACGATTCCTCAATCTCTCCGGA	7225
QY	3181	CAGGAGGTGAGTTCCTGAGTGAAGGCTGTGAATCAAGAAGACCCCTATAGGCTTTGCA	3240
Db	7226	CAGGAGGTGAGTTCCTGAGTGAAGGCTGTGTAAGCTGTGAAGCAAGAAATGCTTATGGCTTGCA	7285
QY	3241	TATGACACCCGCTGTTTGACTCAACAGTCACTGAGAATGACATCGGTGTAAGAGACTCA	3300
Db	7286	TATGACACCCGCTGTTTGACTCAACAGTCACTGAGAATGACATCCGTGTTGAGAGACTCA	7345
QY	3301	ATTATATCAATGTGTGACTTGGCCCCCGAAGCCAGACAGGCCATTAAGTGGCTCACAG	3360
Db	7346	ATTATACCAATGTGTGACTTGGCCCCCGAAGCCAGACAGGCCATTAAGTGGCTCACAG	7405
QY	3361	CGGCTTTATATCGGAGGTCCCTCTGACTTAATTCAAAAGGCGAGACTCGGCTATCGCGG	3420
Db	7406	CGGCTTTATATCGGAGGGGCCCTCTGACTTAATTCATAAGGCGAGACTCGGCTATCGCGG	7465
QY	3421	TGCGCGCGAGCGGCGTGTGACGACTGAGCTGCGGTATATCCCTCACATGTTACTTGAAG	3480
Db	7466	TGCGCGCGAGCGGCTGTGACTGACCAACAGCTGCGGTATATCCCTCACATGTTACTTGAAG	7525
QY	3481	GCCCTGTGAGCTGTGAGGCTGCAAAAGCTCCAGGACTGCGACGATGCTCGTGTGCGAGAC	3540
Db	7526	GCCCTGTGAGCTGTGAGGCTGTGAGGCTGCGAAGCTCCAGGACTGCTGATGTGCGAGAC	7585
QY	3541	GGCCTTGTGCTTATCTGTGAGAGCGCGGAAACCCAGAGAGACGCGCGAGCTTACGAGTC	3600
Db	7586	GACCTTGTGCTTATCTGTGAAAGCGCGGAAACCAAGAGACGAGCGGAGCTTACGAGGTC	7645
QY	3601	TTTCAACGAGGCTATAGACTAGGTAATCTTGCCCCCGGGGAGCCCGCCCAACCAAGATAC	3660
Db	7646	TTTCAACGAGGCTATAGACTAGGTAATCTTGCCCCCGGGGAGCCCGCCCAACCAAGATAC	7705
QY	3661	GACCTGAGTGTATACATGATGCTCTCCATATGTGTGCGTGTGCGGACGATGATCTTGGC	3720
Db	7706	GACTTGAAGTGTATACATGATGCTCTCCATATGTGTCAATTCGCGGACGATGATCTTGGC	7765
QY	3721	AAAAAGGTATATCTACCTCACCCGTGACCC	3749
Db	7766	AAAAAGGTATATCTATCTCACCCGTGACCC	7794

RESULT 13  
US-10-029-907-4  
Sequence 4, Application US/10029907  
Patent No. 6706874  
GENERAL INFORMATION:  
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
TITLE OF INVENTION: HEPATITIS C VIRUS  
FILE REFERENCE: 13/083  
CURRENT APPLICATION NUMBER: US/10/029,907  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/257,857  
PRIOR FILING DATE: 2000-12-22  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 8643  
TYPE: DNA  
ORGANISM: HCV  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1802) ... (8407)  
US-10-029-907-4

Query Match 87.7%; Score 3288.2; DB 4; Length 8643;  
Best Local Similarity 92.3%; Pred. No. 0;  
Matches 3461; Conservative 0; Mismatches 288; Indels 0; Gaps 0;

1 TGGAGAGGCGCTTCAAGAGCTCAAGGAGTGGATGAGCCACTTCCTGTCGCAAGCAAG 60  
4046 TGGAGAGGCGCTTCAAGAGCTCAAGGAGTGGATGAGCCACTTCCTGTCGCAAGCAAG 4105  
61 CAGGCGAGAGCAACTTCCCTTACCTGAGTGGCGTACCAAGCTACTGTGTGCTTAAGGCC 120  
4106 CAGGCGAGAGCAACTTCCCTTACCTGAGTGGCGTACCAAGCTACTGTGTGCTTAAGGCC 4165  
121 CAGGCGCGGCGCTTCAAGAGCTCAAGGAGTGGATGAGCCACTTCCTGTCGCAAGCAAG 180  
4166 CAGGCGCGGCGCTTCAAGAGCTCAAGGAGTGGATGAGCCACTTCCTGTCGCAAGCAAG 4225  
181 CTGCGCGGCGCTTCAAGAGCTCAAGGAGTGGATGAGCCACTTCCTGTCGCAAGCAAG 240  
4226 CTGCGCGGCGCTTCAAGAGCTCAAGGAGTGGATGAGCCACTTCCTGTCGCAAGCAAG 4285  
241 ACAACCCCTTAACCAATTATCATATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 300  
4286 ACAACCCCTTAACCAATTATCATATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 4345  
301 AGGACCTGGTGTGAGTGGAGGAGTCTTGAAGCTTGGCTGATGATGATGATGATGATGAT 360  
4346 AGGACCTGGTGTGAGTGGAGGAGTCTTGAAGCTTGGCTGATGATGATGATGATGATGAT 4405  
361 GGGAGCGTGTCTATGAGTGGAGTATGATGATGATGATGATGATGATGATGATGATGAT 420  
4406 GGGAGCGTGTCTATGAGTGGAGTATGATGATGATGATGATGATGATGATGATGATGAT 4465  
421 AGGGAATGCTCTTACCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 480  
4466 AGGGAATGCTCTTACCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 4525  
481 ATGAGAGGAGGAAATGAGAGTGGAGGAGTCTTGAAGCTTGGCTGATGATGATGATGATGAT 540  
4526 ATGAGAGGAGGAAATGAGAGTGGAGGAGTCTTGAAGCTTGGCTGATGATGATGATGATGAT 4585  
541 ACAGCCCAACCAAGAGGAGGAGGAGTCTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 600  
4586 ACAGCCCAACCAAGAGGAGGAGGAGTCTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4645  
601 GAGACCTTCTGGGGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
4646 GAGACCTTCTGGGGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4705

661 TTGTCCACTCTGCTGGAGATCCCGGATTTGATCATGATGAGCGCTTCAAGACCTCTGTG 720  
4706 TTGTCCACTCTGCTGGAGATCCCGGATTTGATCATGATGAGCGCTTCAAGACCTCTGTG 4765  
721 ACTAGCCGCTTACCAACCAATCTACCTCTCTGCTTAACTCTGAGGAGATGAGGATGAG 780  
4766 ACTAGCCGCTTACCAACCAATCTACCTCTCTGCTTAACTCTGAGGAGATGAGGATGAG 4825  
781 GCCCAACTGCTCCCCGAGTGTGCTGCTTCAAGCTTGTGATGAGGAGGAGGAGGAGGAG 840  
4826 GCCCAACTGCTCCCCGAGTGTGCTGCTTCAAGCTTGTGATGAGGAGGAGGAGGAGGAG 4885  
841 GCTGTGGAGCATAGGCTTTGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 900  
4886 GCTGTGGAGCATAGGCTTTGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 4945  
901 GGAATGAGAGGCGGCTGCTGCTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 960  
4946 GGAATGAGAGGCGGCTGCTGCTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 5005  
961 GACCTGTAACTTACTCCCTGCAATCCCTCTCTGCTGAGTGGAGGAGGAGGAGGAGGAG 1020  
5006 GACCTGTAACTTACTCCCTGCAATCCCTCTCTGCTGAGTGGAGGAGGAGGAGGAGGAG 5065  
1021 TGGAGAGCATATGCTGAGTGGAGAGTGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
5066 TGGAGAGCATATGCTGAGTGGAGAGTGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5125  
1081 CGGCTGATAGCTTGGCTTGGCGGGGTAACATGTTTCCCTCCCAAGCATATGATGATGATGAT 1140  
5126 CGGCTGATAGCTTGGCTTGGCGGGGTAACATGTTTCCCTCCCAAGCATATGATGATGATGAT 5185  
1141 AGGAGCGCGAGACAGTGTCACTGAGATCTCTCCGACCTTACTACTACCAACCAAGCTG 1200  
5186 AGGAGCGCGAGACAGTGTCACTGAGATCTCTCCGACCTTACTACTACCAACCAAGCTG 5245  
1201 AAGAGCTTCAACAGTGAATTAACAGAGCTGCTCCAGCCTGCTCGGCTGTGAGTGA 1260  
5246 AAGAGCTTCAACAGTGAATTAACAGAGCTGCTCCAGCCTGCTCGGCTGTGAGTGA 5305  
1261 AAGAGTGTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
5306 AAGAGTGTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5365  
1321 AAGCTCTGCGGAGATTAACGAGGAGTCCCTTTTCTCATGCAACAGTGGATGAACAGGAG 1380  
5366 AAGCTCTGCGGAGATTAACGAGGAGTCCCTTTTCTCATGCAACAGTGGATGAACAGGAG 5425  
1381 GTCTGCGGAGAGACGAGCATATGACAGCACCTGCTCATGTGAGACAGATCACCGGA 1440  
5426 GTCTGCGGAGAGACGAGCATATGACAGCACCTGCTCATGTGAGACAGATCACCGGA 5485  
1441 CATGTCAAAAACGCTTCCATGAGAGTGTGAGGCTTAAAGCTGTATGATTAACATGTGAGCAT 1500  
5486 CATGTCAAAAACGCTTCCATGAGAGTGTGAGGCTTAAAGCTGTATGATTAACATGTGAGCAT 5545  
1501 GGAACATTCCTCCATCAACGATTAACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
5546 GGAACATTCCTCCATCAACGATTAACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5605  
1561 TATTCAGAGGAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1620  
5606 TATTCAGAGGAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5665  
1621 GATTTCACTAGCTGACAGCATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1680  
5666 GATTTCACTAGCTGACAGCATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 5725  
1681 GCGCCGGAATTCCTTCAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740  
5726 GCGCCGGAATTCCTTCAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5785  
1741 AAACCTCTCTTCAAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1800



Db	5786	AAACCCCTCCTAACGGAGAGAGGTACATTCTCGGTGGGGCTCTAATCATTA	CTGGTTGGG	5845
QY	1801	TCGCAGCTTCCATGTGAGCCCGAACCAGATGTAGCAGTCTOACTTCCATGCTCA	CCGAC	1860
Db	5846	TCACAGCTCCCATGTGAGACCCGAACTGGACGTAGCAGTCACTCACTTCATCTCA	CCGAC	5905
QY	1861	CCCTCCCAATCAACAGCAGAGACGGCTTAAAGGCGAGGCTGGCCAGGGGGTCT	CCCCCTCC	1920
QY	1921	TTGGCAGACTTTCAGTACAGCCAGTGTGTGCGCTTCTCGAAGGCGACATTA	CATTACC	1980
Db	5966	TTGGCAGACTCATCAGCTAAGCCAGTGTCTGTGGCTTCTTGAGGCAATGCA	CTAAC	6025
QY	1981	CAAAATGACTTCCAGACGCTTGACCTCATGAGGCAACCTTGTGTGGCGCAT	GAGATG	2040
Db	6026	CGTATGACTCCCGGAGCGTGAACCTCATGAGGCGCAACTCTGTGGCGCAG	AGATG	6085
QY	2041	GGCGGGGACATTACCCCGCTGGAAGTCAGAGAAACAAGATATATCTTGACT	CTTTGGAC	2100
Db	6086	GGCGGGAAATCAACCCGGGTGAAGTCAAAAAATAGATATTTTGGACTTTTG	AG	6145
QY	2101	CCGCTCCGAGCGGAGAGAGATGAGCGGGAGATGTCGTCGCGCGGAGATCT	CGGGAAA	2160
Db	6146	CCGCTCCAGCGGGAGAGAGATGAGAGGAAATATCCGTTCCGGCGGAGATCT	CGGGAG	6205
QY	2161	TCCAAAATTTCCACACAGCATGCGCCGATGAGGACCGCCGATTAACA	CCCTCCGCTG	2220
Db	6206	TCCAAAATTTCCCTCGAGCGATGCGCATATGAGGACGCCGGATTAACA	CCCTCACTG	6265
QY	2221	CTGAGATCTTGAGAGGCCCCCGGACTAGTCCCTCCATGTGTATCATGGT	GCCACTGCA	2280
Db	6266	TTGGAGTCTTGAGAGGACCCCGGACTAGTCCCTCCAGTGTATCAACGG	GTGTCCATTGCGG	6325
QY	2281	CCCTATAGACCCCTCCTATACACCTCCACGAGAGAAAGAGACAGTTGTCT	ATACGAA	2340
Db	6326	CCCTCGAAGGCCCTCCGATACACCTCCACGAGAGAAAGAGACGATGTCT	GTACGAA	6385
QY	2341	TCACCGTGTCTTCTGCCCCCTGGCGAGCTTCCCAAAAGGCTTTTCGAT	CTCCGAAACG	2400
Db	6386	TCACCGTGTCTTCTGCCCCCTGGCGAGCTTCCCAAAAGGCTTTCGAT	CTCCGAAATG	6445
QY	2401	TCGGCCGTGACACGCGGACCGGCAACGCCCTCTCTGACCAACCTTCG	ACGACGGCGGA	2460
Db	6446	TCGGCCGTGACACGCGGACCGGCAACGCCCTCTCTGACACGCCCTCG	ACGACGGGAC	6505
QY	2461	GCAGGATCTGACGTTGAGTCCGATTTCTCTCCATGCCCCCTTGAGAGG	GGAGACCGGGGAC	2520
Db	6506	GCAGGATCTGACGTTGAGTCCGATTTCTCTCCATGCCCCCTTGAGAGG	GGAGACCGGGGAT	6565
QY	2521	CCCGATCTCAGGACCGGCTCTTGGCTTACCGTAGTGAAGGCGCGGTGA	GAGAGTCTGCTC	2580
Db	6566	CCCGATCTCAGGACCGGCTCTTGGCTTACCGTAGTGAAGGCGCGGTG	AGAGTCTGCTC	6625
QY	2581	TGCTGCTGATGTCTTACATGACACAGGCGCTCTGATCAGCGCATGTG	CTCGGAGAA	2640
Db	6626	TGCTGCTGATGTCTTACATGACACAGGCGCGCTGATCAGCGCATGTG	CTCGGAGAA	6685
QY	2641	AGCAAGCTGCCCCATCAACGCGTTGAGCACTTTTGTGCTCACACACA	TAGTGTAC	2700
Db	6686	AGCAAGCTGCCCCATCAACGCGTTGAGCACTTTTGTGCTCACACACA	TAGTGTAT	6745
QY	2701	GCTACCAATCCCGGAGCGCAAGCCAGGCGGAGAAAGTCACTTTGA	CAGACTGCA	2760
Db	6746	GCTACCAATCTCGACGCGCAAGCTTGGCGGAGAAAGTCACTTTGA	CAGACTGAG	6805
QY	2761	ATCTCGACGATCACTACAGAGCGTGTCTCAAGAGATGAAGGCGAAG	CGCTTCCAGATT	2820
Db	6806	GTCTCGACGACCACTACCGGAGCGTGTCTCAAGAGATGAAGGCGAAG	CGCTTCCAGATT	6865
QY	2821	AAGGTAAGCTTTATCAGTGAAGAGCTTGCAGACTGACGCCCTCCACAT	TTGGCCCAA	2880

Db	6866	AAGGCTAAACCTTATCCGTCCGTGAGGAAGACCTGTAAAGTGAAGCCGCCCAANTTGGCCAGA	6925
Qy	2881	TCCTAAATTTGGCTATGGGGCAAAAGAGAGTCCGGAAACCTATTCACAGACAGCCATTAAACAC	2940
Db	6926	TCCTAAATTTGGCTATGGGGCAAAAGAGAGTCCGGAAACCTATTCACAGACAGCCATTAAACAC	6985
Qy	2941	ATCCGCTCCGGTGGGAGACCTTGTGTGAAGAACACTGAACAACCAATTGACACCAACATC	3000
Db	6986	ATCCGCTCCGGTGGGAGACCTTGTGTGAAGAACACTGAACAACCAATTGACACCAACATC	7045
Qy	3001	ATGGCAAAAATGATGAGTTTTCTGCGTCCAAACAGAGAGAGAGGCGCAAGCCAGCTCGC	3060
Db	7046	ATGGCAAAAATGATGAGTTTTCTGCGTCCAAACAGAGAGAGAGGCGCGCAAGCCAGCTCGC	7105
Qy	3061	CTTATTCGTGTCCCAAGCTTGGGGGTCCGTGTGTGCGAAGAAATGCGCTTATGACGTG	3120
Db	7106	CTTATTCGTATTTCCAGATTGGGGGGTTCGTGTGTGCGAAGAAATGCGCTTATGACGTG	7165
Qy	3121	GTCACCAACCCCTCCAGGCGTGTGATGGGGCTCCGTGATCGGAATTCAGATATTCCTCGGA	3180
Db	7166	GTCACCAACCCCTCCAGGCGTGTGATGGGGCTCCGTGATCGGAATTCAGATATTCCTCGGA	7225
Qy	3181	CAGCGGGTCGAGTTCCTCGTGTGAAGCGCTGGAAATCAAGAAGAACCCCTATGSGCTTTGCA	3240
Db	7226	CAGCGGGTCGAGTTCCTCGTGTGAAGCGCTGGAAAGCGGAAGAAATGCCCTATGSGCTTGCA	7285
Qy	3241	TATGACACCCGCTGTTTGAAGTCAACAGTCACTGAGATGACATCCGTGTGAGAGAGTCA	3300
Db	7286	TATGACACCCGCTGTTTGAAGTCAACAGTCACTGAGATGACATCCGTGTGAGAGAGTCA	7345
Qy	3301	ATTATATCAATGTTGTGACTTGGCCCCCGAAGCCAGAGAGGCCATTAAGTGGCTCACAGAG	3360
Db	7346	ATTATCACAATGTTGTGACTTGGCCCCCGAAGCCAGAGAGGCCATTAAGTGGCTCACAGAG	7405
Qy	3361	CGGCTTATATACGGGGGTGCCCTGTGACTCAATATTCAAAAAGGAGAACTCGGCTATCGCCGG	3420
Db	7406	CGGCTTATATACGGGGGGCCCCGTGACTCAATATTCCTTAAAGGCGAAGCTCGGCTATCGCCGG	7465
Qy	3421	TGCCGCGCAGACGGCGGTGTGACGACGACTGAGCTGAGTAATACCTCACATGTTAATTGAAG	3480
Db	7466	TGCCGCGCAGACGGGTGTGACTGACGACGACTGAGCTGAGTAATACCTCACATGTTAATTGAAG	7525
Qy	3481	GCCCTTCGAGCTGTGCAAGCTGCAGAAAGTCCAGAGCTGCACAGATGCTCGTGTGGGAGAC	3540
Db	7526	GCCGCTCCGCGCTGTGCAAGCTGCAGAAAGTCCAGAGCTGCACAGATGCTCGATGTGGGAGAC	7585
Qy	3541	GGCGCTTCGCTAATCTGTGAGAGAGCGGAGAACCCAGAGAGACGCGGCGAGCGTACAGAGTC	3600
Db	7586	GACCTTTCCTGTATCTGTGAAAGCGCGGGAGCCCAAGAGAGACGAGCGAGCGCTTACGGGGCC	7645
Qy	3601	TTCAACGAGGCTATGACTAGAGTACTCTGCCCCCGCGGGAGCCCGCCCAACCAAGATAC	3660
Db	7646	TTCAACGAGGCTATGACTAGATTAATCTGCCCCCTGTGGGAGCCCGCCCAACCAAGATAC	7705
Qy	3661	GACCTGAGTGTGATACATATGCTCTCCAAATGTGTGCGTCCGCGACGATGCACTGTGC	3720
Db	7706	GACTTGGAGTGTGATACATATGCTCTCCAAATGTGTGCACTGCGGACGATGCACTGTGC	7765
Qy	3721	AAAAGGCTATATCAATCCACCCGAGACC 3749	
Db	7766	AAAAGGCTATATCAATCCACCCGAGACC 7794	

RESULT 14  
US-10-029-907-5  
; Sequence 5, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; FILE REFERENCE: 13/063  
; CURRENT APPLICATION NUMBER: US/10/029,907

;; CURRENT FILING DATE: 2001-12-21  
;; PRIOR APPLICATION NUMBER: 60/257,857  
;; PRIOR FILING DATE: 2000-12-22  
;; NUMBER OF SEQ ID NOS: 25  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 5  
;; LENGTH: 8648  
;; TYPE: DNA  
;; ORGANISM: HCV  
;; FEATURE:  
;; NAME/KEY: CBS  
;; LOCATION: (1802) ... (8407)  
US-10-029-907-5

Query Match 87.6%; Score 3285; DB 4; Length 8648;  
Best Local Similarity 92.3%; Pred. No. 0;  
Matches 3459; Conservative 0; Mismatches 290; Indels 0; Gaps 0;

QY 1 TGGAGAGGCGCTCTTCAAGGCTCACCACGATGATGCCACTTCTGTCCAAAGAG 60  
DB 4046 TGGAGAGGCGCTCTTCAAGGCTCACCACATAGAGCCCATTTCTGTCCAGACTAG 4105  
QY 61 CAGGCAAGAGACAATCTCCCTTACCTGTGGCGGTACAGGCTACTGTGTGCGCTAGAGCC 120  
DB 4106 CAGGCAAGAGACAATCTCCCTTACCTGTGTAGCATACAGGCTACGGGTGCGCAGAGCT 4165  
QY 121 CAGGCGCCACCTCATCATGATGATCAAAATGTGAAGTGTCTCATACGGCTTAAGCCACT 180  
DB 4166 CAGGCTTCACCTCATCTGTGGAGCAAAATGTGAAGTGTCTCATACGGCTTAAGCCACT 4225  
QY 181 CTGCGGCGGCAACACCTTGTCTGTATAGGCTGGAGGCGCTCCAAAAGAGTCAACCTTC 240  
DB 4226 CTGCGAGGCGCAACGCGCTGTGTATAGGCTGGAGGCGCTTCAAAAGAGTCAACCTTC 4285  
QY 241 ACACACCCCTTAACCAATTCATCATGTGCATGTCAAGCCAGCTGGAGGTGTGTACG 300  
DB 4286 ACACACCCCTTAACCAATTCATCATGTGCATGTCAAGCTGAGCTGAGGTGTGTACG 4345  
QY 301 AGACCTGGGTGTGGTGGGCGGGGTCTTGCAGCTCTGGCTGGTATGTGCTTGAACA 360  
DB 4346 AGACCTGGGTGTGGTGGGCGGGGTCTTGCAGCTCTGGCTGGTATGTGCTTGAACA 4405  
QY 361 GGCAGCTGTGTCTTGTGGTGAAGATCATCTTGTCCGGGCGGCGGCTATTTGTCCGAC 420  
DB 4406 GGCAGCTGTGTCTTGTGGGCAAGATCATCTTGTCCGGAAGCGGCAATTCATTTCCGAC 4465  
QY 421 AGGGAAGTCTCTTACCAAGAGTGTGATGATGGAAGAGTGGCGGTGCACCTTCCCTTAC 480  
DB 4466 AGGGAAGTCTCTTACCGGAGTTCGATGATGGAAGAGTGGCGGTGCACCTTCCCTTAC 4525  
QY 481 ATGAGAGAGGAATGAGCTCGCGAGACATTCAGCAAAAGCGCTCGGGTTGTGACAG 540  
DB 4526 ATGAGAGAGGAATGAGCTCGCGAGACATTCAGCAAAAGCGAATCGGGTTGTGACAA 4585  
QY 541 ACAGCCACAGCAAGCGAGGCGCTGCTCCGCTGTGAGTCAAGTGGCGAGCCCTT 600  
DB 4586 ACAGCCACAGCAAGCGAGGCGCTGCTCCGCTGTGAGTCAAGTGGCGAGCCCTT 4645  
QY 601 GAGACTCTTGTGGCGAACAATGTGAAGCTTCATACGCGGATACAGTACTTAAGCAGGC 660  
DB 4646 GAGACTCTTGTGGCGAACAATGTGAAGTTCATACGCGGATACAGTACTTAAGCAGGC 4705  
QY 661 TTGTCCACTGCTGTGGGAATCCCGGATGATCATGATGTGGCTTCAAGGCTCTGTC 720  
DB 4706 TTGTCCACTGCTGTGGCAACCCCGGATGATCATGATGTGGCTTCAAGGCTCTTATC 4765  
QY 721 ACTAGCCGCTACACCAATCTACCTCTGTCTTAACATCTGTGGGAGTGGGTAGCC 780  
DB 4766 ACTAGCCGCTACACCAATCTACCTCTGTCTTAACATCTGTGGGAGTGGGTAGCC 4825  
QY 781 GCCCAACTGTCTCCCGCAAGTGTCTTGTAGCTTGTGAGGCGCGCAATGTCTGTGCG 840  
DB 4826 GCCCAACTGTCTCCCGCAAGTGTCTTGTAGCTTGTGAGGCGCGCAATGTCTGTGAGCG 4885

QY 841 GCTGTGGCAGATAGGCTTGGGAAGTGTCTGTGACATCTTGGCGGCTATGAGCA 900  
DB 4886 GCTGTGGCAGATAGGCTTGGGAAGTGTCTGTGATATTTTGGCAGGTTATGAGCA 4945  
QY 901 GGAATGAGCAGGCGGCTGTGGCTTTAAGTCAAGAGCGGAATGCGCTCCACCGAG 960  
DB 4946 GGAATGAGCAGGCGGCTGTGGCTTTAAGTCAAGAGCGGAATGCGCTCCACCGAG 5005  
QY 961 GACTGTGTTAATCTACTCCCTGCATACCTCTCTCTGAGGCTGTGGTGTGAGTGG 1020  
DB 5006 GACTGTGTTAATCTACTCCCTGCATACCTCTCTCTGAGGCTGTGGTGTGAGTGG 5065  
QY 1021 TCGCAGCGATACCTGTGTGCGACGTGGTTCAGGGAGGGGCTGTGAGTGAAC 1080  
DB 5066 TCGCAGCGATACCTGTGTGCGACGTGGTTCAGGGAGGGGCTGTGAGTGAAC 5125  
QY 1081 CGGCTGATAGCTTGTGCTGTGGGGTAAACATGTTTCCCGCAGCATATGTGCCAG 1140  
DB 5126 CGGCTGATAGCTTGTGCTGTGGGGTAAACATGTTTCCCGCAGCATATGTGCTGAG 5185  
QY 1141 AGGAGCGCGGAGACGTGTCACTGATCTCTCCGACCTTACTATCAACCCACTGTG 1200  
DB 5186 AGGAGCGCTGACAGACGTGTCACTGATCTCTCTCTGATCTTACTATCACTAGCTGTG 5245  
QY 1201 AAGAGGCTCCACAGTGTATTAACGAGACTGCTCCACGCGCTGTCTCCGGCTGTGACT 1260  
DB 5246 AAGAGGCTTCACAGTGTATTAACGAGACTGCTCCACGCGCTGTCTCCGGCTGTGACT 5305  
QY 1261 AAGGATGTTTGGGACTGATATGACAGTGTGCTGACTTCAAGACCTGTCTCCAGTCC 1320  
DB 5306 AAGGATGTTTGGGATGATATGACAGTGTGCTGACTTCAAGACCTGTCTCCAGTCC 5365  
QY 1321 AAGCTCTGCGCGGATTAACGGGAGTCCCTTTTCTCATGCAAGTGGATCAAGGGG 1380  
DB 5366 AAGCTCTGCGCGGATTAACGGGAGTCCCTTTTCTCATGCTGATGACAGTGGATCAAGGG 5425  
QY 5426 GTCTGGCGGGGAGCGGATCATGCAACCACTGCTCCATGTGAGCAAGATCACCGGA 5485  
QY 1441 CATGTCAAAAACGCTTCATGAGAGTGTGGGCTTAAGAAGCTGTGATGATCATGTGGCAT 1500  
DB 5486 CATGTCAAAAACGCTTCATGAGAGTGTGGGCTTAAGAAGCTGTGATGATCATGTGGCAT 5545  
QY 1501 GGAACATTCGCCATCAAGCATACCAAGGCGCTGTGACGCGCTCCCAAGCGCCAAC 1560  
DB 5546 GGAACATTCGCCATTAACGCTACACAGGCGCTGTGACGCGCTCCCAAGCGCCAAC 5605  
QY 1561 TATTCAGGGCGCTGTGGGCTGTGCTGTGAGAGTACGTGAGGTTACGGGCTGAGG 1620  
DB 5606 TATTCAGGGCGCTGTGGGCTGTGCTGTGAGAGTACGTGAGGTTACGGGCTGAGG 5665  
QY 1621 GATTTCACTACGTGACAGCATGACCACTGACCAAGTAAATGCCCGTCCAGGTTCCA 1680  
DB 5666 GATTTCACTACGTGACAGCATGACCACTGACCAAGTAAATGCCCGTCCAGGTTCCA 5725  
QY 1681 GCGCCGGAATTTCTCAAGAAGTGAAGGAGTGGCGGCTGACAGATAGCGCTCGGGTGC 1740  
DB 5726 GCGCCGGAATTTCTCAAGAAGTGAAGGAGTGGCGGCTGACAGATAGCGCTCGGGTGC 5785  
QY 1741 AAACCTCTCTACGAGAGAGGTCATATTCAGAGTGGGCTCAACCAATACCTGTGTGG 1800  
DB 5786 AAACCTCTCTACGAGAGAGGTCATATTCAGAGTGGGCTCAACCAATACCTGTGTGG 5845  
QY 1801 TCGAGCTCCATGCGAGCGCGAACCAGATGTAGCAGTGTCTCATGTCTCACCGAC 1860  
DB 5846 TCGAGCTCCATGCGAGCGCGAACCAGATGTAGCAGTGTCTCATGTCTCACCGAC 5905  
QY 1861 CCTTCCCAATACAGCAGAGAGGTTAAGCGAGCTGTGCGAGGAGGTCTCCCGCTCC 1920  
DB 5906 CCTTCCCAATACAGCAGAGAGGTTAAGCGAGCTGTGCGAGGAGGTCTCCCGCTCC 5965



QY 1921 TTGGCAGCTCTTACAGTACGAGTGTCTGCGCTTCTCGAAGGCACTACATTACC 1980  
DB 5966 TTGGCAGCTCTTACAGTACGAGTGTCTGCGCTTCTCGAAGGCACTACATTACC 6025  
QY 1981 CAAATGACTTCCAGAGCTGACCTGATGAGGCCAATCTCTGTGGCGGCAATGAGTGT 2040  
DB 6026 CGTACGACTCTCCCGAGCGTACCTGATGAGGCCAATCTCTGTGGCGGCAATGAGTGT 6085  
QY 2041 GGGGGGAGCAATTAACCGCGTGAAGTGAAGAAAGGATGATATCCGTGACTCTTTCGAC 2100  
DB 6086 GGGGGGAGCAATTAACCGCGTGAAGTGAAGAAAGGATGATATCCGTGACTCTTTCGAC 6145  
QY 2101 CGGCTCCGAGCGAGAGAGTGAAGCGGAAAGTGTCCGCTCCGCGGAGATCTTCCGAA 2160  
DB 6146 CGGCTCCGAGCGAGAGAGTGAAGCGGAAAGTGTCCGCTCCGCGGAGATCTTCCGAA 6205  
QY 2161 TCCAGAAATTTCCACAGCGATGCGCGGATGCGGCAAGCGGATTAACAACCTTCCGCTG 2220  
DB 6206 TCCAGAAATTTCCACAGCGATGCGCGGATGCGGCAAGCGGATTAACAACCTTCCGCTG 6265  
QY 2221 CTGAGACTCTGGAAGGCGCGGACTAGTCCCTCGATGCTGATGAGGCGGCACTGCA 2280  
DB 6266 TTGAGACTCTGGAAGGCGCGGACTAGTCCCTCGATGCTGATGAGGCGGCACTGCA 6325  
QY 2281 CCTACTAGAACCCCTCTCTATACCACTCCAGAGAGAGAGAGAGAGTGTCTGAGAGAA 2340  
DB 6326 CTGCGCAAGGCGCGCTCGATACCACTCCAGAGAGAGAGAGAGTGTCTGAGAGAA 6385  
QY 2341 TCCACCGTGTCTTCTGCGTGGGAGAGCTTGGCAAAAGCTTTCGATGCTCGAAGCG 2400  
DB 6386 TCCACCGTGTCTTCTGCGTGGGAGAGCTTGGCAAAAGCTTTCGATGCTCGAAGCG 6445  
QY 2401 TGGCGCGTGAAGCGGAG 2460  
DB 6446 TGGCGCGTGAAGCGGAG 6505  
QY 2461 GCAGAGTCTGACCTGAGTGTCTTCTCTCAATGCGCGCTTGAAGGAGAGAGAGAGAGAG 2520  
DB 6506 GCGGAGTCTGACCTGAGTGTCTTCTCTCAATGCGCGCTTGAAGGAGAGAGAGAGAGAG 6565  
QY 2521 CCGGATCTGACGAGCGGAGTGTCTTCTCTCAATGCGCGCTTGAAGGAGAGAGAGAGAG 2580  
DB 6566 CCGGATCTGACGAGCGGAGTGTCTTCTCTCAATGCGCGCTTGAAGGAGAGAGAGAGAG 6625  
QY 2581 TGTGTCTGATGTCTTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
DB 6626 TGTGTCTGATGTCTTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6685  
QY 2641 AGCAAGCTGCGCATCAAGCGGTTGAGCAACTTTGTGCGGTGACCAACATGCTGAC 2700  
DB 6686 AGCAAGCTGCGCATCAAGCGGTTGAGCAACTTTGTGCGGTGACCAACATGCTGAC 6745  
QY 2701 GCTACCAATCCGAGCGGAG 2760  
DB 6746 GCTACCAATCCGAGCGGAG 6805  
QY 2761 ATCTTGAAGAGATCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
DB 6806 ATCTTGAAGAGATCACTACAG 6865  
QY 2821 AAGGCTAGCTTCTATCAGTAG 2880  
DB 6866 AAGGCTAGCTTCTATCAGTAG 6925  
QY 2881 TCTAAATTTGGCTATGAGGAG 2940  
DB 6926 TCTAAATTTGGCTATGAGGAG 6985  
QY 2941 ATCCGCTCCGTGTGGAG 3000  
DB 6986 ATCCGCTCCGTGTGGAG 7045  
QY 3001 ATGGCAAAAATGAGGTTTTTCTGCTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060

DB 7046 ATGGCAAAAATGAGGTTTTTCTGCTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7105  
QY 3061 CTTATGATGTTCCCAACTTGGGGGTCGTTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120  
DB 7106 CTTATGATGTTCCCAACTTGGGGGTCGTTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7165  
QY 3121 GTCTCAACCTCCCTCAGAGCTGATGAGGCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 3180  
DB 7166 GTCTCAACCTCCCTCAGAGCTGATGAGGCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 7225  
QY 3181 CAGCGGCTGAGTCTCTGCTGAG 3240  
DB 7226 CAGCGGCTGAGTCTCTGCTGAG 7285  
QY 3241 TATGAGACCCGCTGTTTGAATCAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
DB 7286 TATGAGACCCGCTGTTTGAATCAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7345  
QY 3301 ATTATCAATGTTGATCTTGGGCGGAG 3360  
DB 7346 ATTATCAATGTTGATCTTGGGCGGAG 7405  
QY 3361 CCGCTTATATCGGGGTCCTCTGATTAATCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420  
DB 7406 CCGCTTATATCGGGGTCCTCTGATTAATCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7465  
QY 3421 TGGCGGCGAG 3480  
DB 7466 TGGCGGCGAG 7525  
QY 3481 GCGCTGAG 3540  
DB 7526 GCGCTGAG 7585  
QY 3541 GCGCTGAG 3600  
DB 7586 GCGCTGAG 7645  
QY 3601 TTACAG 3660  
DB 7646 TTACAG 7705  
QY 3661 GACCTGAG 3720  
DB 7706 GACCTGAG 7765  
QY 3721 AAAAGGATTAATCACTCAACCGTGAACC 3749  
DB 7766 AAAAGGATTAATCACTCAACCGTGAACC 7794

RESULT 15  
US-09-539-601-22  
; Sequence 22, Application US/09539601C  
; Patient No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenischlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; EARLIER FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 8001  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5' UTR  
; LOCATION: (1) .. (341)



Db 5209 TCACAGCTCCCAAGCGAGCCGAAACGAGTAGTAGAGTGTCTCACTTCATGCTCACGAC 5268  
 Qy 1861 CCTTCCCAATACACAGCAGAGAGCGCTAAGCGAGCTGGCCAGGGGGTCTCCCTCCCTC 1920  
 Db 5269 CCTTCCCAATACAGCAGAGAGCGCTAAGCGAGCTGGCCAGGGGGTCTCCCTCCCTC 5328  
 Qy 1921 TTGGCAGCTCTTCACTAGCAGCTAGTGTCTGCGCTTCTCTGAGGCGACATACATTACC 1980  
 Db 5329 TTGGCAGCTCTTCACTAGCAGCTAGTGTCTGCGCTTCTCTGAGGCGACATACATTACC 5388  
 Qy 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGGCGGACATGAGT 2040  
 Db 5389 CTTATGACTTCCCGAGCGCTGACCTCATGAGGCCAACCTCTGTGGCGGACATGAGT 5448  
 Qy 2041 GCGCGGAGCATTAACCGCGCTGAGTCAAGAGACAAAGTAGTAATCTCTGACTTTTCAC 2100  
 Db 5449 GCGCGGAGCATTAACCGCGCTGAGTCAAGAGACAAAGTAGTAATTTTGGACTCTTTTCAG 5508  
 Qy 2101 CCGCTCCAGCGGAGGAGTAGAGCGGAAAGTGTCCGTCGCGGCGAGATCTTGCGGAAA 2160  
 Db 5509 CCGCTCCAGCGGAGGAGTAGAGCGGAAAGTGTCCGTCGCGGCGAGATCTTGCGGAG 5568  
 Qy 2161 TCCAAAGAAATTCACACAGCGATGCGCGATGCGCAAGCGGATTACAACTCTCCGCTG 2220  
 Db 5569 TCCAGAAATTCCTCGAGGAGTCCCATATGCGGACGCGCGATTACAACTCTCCACTG 5628  
 Qy 2221 CTGAGATCTCGAAGGCGCGGACTACGCTCTCTCACTGAGTACATGAGTCCCACTGCA 2280  
 Db 5629 TTAGATCTCTGAGAGAGCCGGACTACGCTCTCACTGAGTACATGAGTCTCATTTGCCG 5688  
 Qy 2281 CCTACTAAGACCCCTCTATACACCTCCACGAGAGAGAGCAGTTGTTCTGACAGA 2340  
 Db 5689 CTGCTCAAGCGCCCTCTCATACCACTTCAAGAGAGAGAGCGGTTGCTCTGACAGA 5748  
 Qy 2341 TCCACCGTGTCTTCTGCGCGAGCTTGCACAAAGGCTTTCGATGCTCGAACCG 2400  
 Db 5749 TCTACCGTGTCTTCTGCGCGAGCTTGCACAAAGGCTTTCGAGCTCGAATG 5808  
 Qy 2401 TCGGCGCTGACAGCGGCGACCGGCAACCGCCCTCTGACCAACCTTCCGACGCGCGA 2460  
 Db 5809 TCGGCGCTGACAGCGGCGACCGGCAACCGGCTCTCTGACGAGCCCTCCGACGCGCGAC 5868  
 Qy 2461 GAGAGATCTGAGCTGAGTGTATCTCTCATGCGCCCTTGAAGGGGAGCGCGGGGAC 2520  
 Db 5869 GCGGAGATCCGAGCTGAGTGTATCTCTCATGCGCCCTTGAAGGGGAGCGCGGGGAT 5928  
 Qy 2521 CCGGATCTGAGCGAGCGGCTTGTGTCTACCGTGAAGGAGCGCGGTGAGGACGTCGTC 2580  
 Db 5929 CCGGATCTGAGCGAGCGGCTTGTGTCTACCGTGAAGGAGCGGTGAGGACGTCGTC 5988  
 Qy 2581 TGTGCTGATGTCTTACATGAGAGAGCGGCTTGTATCAAGCGCATGCGTCGCGAGAA 2640  
 Db 5989 TGTGCTGATGTCTTACATGAGAGAGCGGCTTGTATCAAGCGCATGCGTCGCGAGAA 6048  
 Qy 2641 AAGAGCTGCGATCAACCGCTTGAAGCACTCTTGTCTGTGACACAACTGCTAC 2700  
 Db 6049 AAGAGCTGCGATCAACCGCTTGAAGCACTCTTGTCTGTGACACAACTGCTAC 6108  
 Qy 2701 GGTACACATCCGCGAGCGCAAGCGCGGAGAGAAAGTCACTTTGACAGATGCA 2760  
 Db 6109 GGTACACATCCGCGAGCGCAAGCGCGGAGAGAAAGTCACTTTGACAGATGCA 6168  
 Qy 2761 ATCTGAGAGATCACTACAGAGAGCTGTCTAAGAGATGAAGGCGAGCGTCCACAGTT 2820  
 Db 6169 GTCTGAGAGATCACTACAGAGAGCTGTCTAAGAGATGAAGGCGAGCGTCCACAGTT 6228  
 Qy 2821 AAGGCTAAGCTTCTATCAGTGAAGAGAGCTGTCAAGCTGACGCGCCCACTTTCGCGCAA 2880  
 Db 6229 AAGGCTAAGCTTCTATCAGTGAAGAGAGCTGTCAAGCTGACGCGCCCACTTTCGCGCAA 6288  
 Qy 2881 TCTAATTTGGCTATGGGCGAAAGAGCTCCGGAACCTATCCAGCAAGGCATTAAACAC 2940

Db 6289 TCTAATTTGGCTATGGGCGAAAGAGAGCTCCGGAACCTATCCAGCAAGCGCTTAACAC 6348  
 Qy 2941 ATCCGCTCGGTGGAGAGACTTGTGAGAGACCTGAACACCAATTGACACCACTAC 3000  
 Db 6349 ATCCGCTCGGTGGAGAGACTTGTGAGAGACCTGAACACCAATTGACACCACTAC 6408  
 Qy 3001 ATGGCAAAAATGAGGTTTCTGCGTCAACAGAGAGAGAGCGCGCAAGCGAGCTGCG 3060  
 Db 6409 ATGGCAAAAATGAGGTTTCTGCGTCAACAGAGAGAGAGCGCGCAAGCGAGCTGCG 6468  
 Qy 3061 CTATGCTGTCCAGACTTGGGCGGTGCTGTGTGCGAGAAAATGCGCTTATGACGTG 3120  
 Db 6469 CTATGCTGTCCAGACTTGGGCGGTGCTGTGTGCGAGAAAATGCGCTTATGACGTG 6528  
 Qy 3121 GTCTCAGCCCTCCAGCTGAGTGTGCTCTGTGAGTTCAGATTCCTCTGGA 3180  
 Db 6529 GTCTCAGCCCTCCAGCTGAGTGTGCTCTGTGAGTTCAGATTCCTCTGGA 6588  
 Qy 3181 CAGCGGCTGAGTTCCTGTGAGAGCTGGAATCAAGAGAGAGCCCTATGCGCTTTCGA 3240  
 Db 6589 CAGCGGCTGAGTTCCTGTGAGAGCTGGAATCAAGAGAGAGCCCTATGCGCTTTCGA 6648  
 Qy 3241 TATGACACCGCTGTTTGACTCAACAGTCACTGAGATGACATCTGTGTAGAGATCA 3300  
 Db 6649 TATGACACCGCTGTTTGACTCAACAGTCACTGAGATGACATCTGTGTAGAGATCA 6708  
 Qy 3301 ATTTATCATGTTTGTGACTTGGCCCCCGAAGCGACAGCCATTAAGTCTGACAGAG 3360  
 Db 6709 ATTTATCATGTTTGTGACTTGGCCCCCGAAGCGACAGCCATTAAGTCTGACAGAG 6768  
 Qy 3361 CGGCTTATATGCGGGGCTCCCTGACTAATTAAGAGGCGAGAACTGCGGCTATCGCGG 3420  
 Db 6769 CGGCTTATATGCGGGGCTCCCTGACTAATTAAGAGGCGAGAACTGCGGCTATCGCGG 6828  
 Qy 3421 TGCCTGCGAGCGCGGTGTGACGACTAGCTGCGGTATATCCCTACATGTTACTTGAAG 3480  
 Db 6829 TGCCTGCGAGCGCGGTGTGACGACTAGCTGCGGTATATCCCTACATGTTACTTGAAG 6888  
 Qy 3481 GCTCTGACGCTGTGAGCTGCAAGCTCCAGAGACTGACAGATGCTGTCGCGAGAG 3540  
 Db 6889 GCTCTGACGCTGTGAGCTGCAAGCTCCAGAGACTGACAGATGCTGTCGCGAGAG 6948  
 Qy 3541 GCGCTGTGCTTATCTGTGAGAGCGCGGAAACCAAGAGAGCGCGGAGCTTACGAGTGC 3600  
 Db 6949 GCGCTGTGCTTATCTGTGAGAGCGCGGAAACCAAGAGAGCGCGGAGCTTACGAGTGC 7008  
 Qy 3601 TTACAGAGAGCTATGACTAGTACTTGTCCCGCGCGGAGCGCGCCCAACAGAAATAC 3660  
 Db 7009 TTACAGAGAGCTATGACTAGTACTTGTCCCGCGCGGAGCGCGCCCAACAGAAATAC 7068  
 Qy 3661 GACCTGAGATGATTAACATATGCTCTTCAATGTGTGCTGCGGACGATGCTTGTGC 3720  
 Db 7069 GACCTGAGATGATTAACATATGCTCTTCAATGTGTGCTGCGGACGATGCTTGTGC 7128  
 Qy 3721 AAAAGGATTAACCTACCTGACCGGTGACCC 3749  
 Db 7129 AAAAGGATTAACCTACCTGACCCGTAACC 7157

Search completed: February 21, 2005, 18:58:28  
 Job time : 596 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 10:29:03 ; Search time 11361 Seconds  
(without alignments)  
12564.124 Million cell updates/sec

Title: US-09-664-363-22

Perfect score: 3750

Sequence: 1 TGGAGAGGGGCTTTCACAGG.....ACTACCTCACCCGTCACCG 3750

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapept 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.2	1.3	925	CNS0091P	AL053013 Drosophila
2	48	1.3	552	CF824031	CF824031 EST701413
3	48	1.3	856	CF824624	CF824624 EST702006
4	47.6	1.3	925	CNS0091P	AL053013 Drosophila
5	47.6	1.3	978	CF824954	CF824954 EST702336
6	47	1.3	635	BX905657	BX905657 Leishmani
7	47	1.3	635	LBAR018006	BX54484 Leishmani
8	47	1.3	855	CO015154	CO015154 EST78536
9	47	1.3	909	CF823313	CF823313 EST700695
10	47	1.3	914	CF826948	CF826948 EST704330
11	47	1.3	923	CF823378	CF823378 EST700760
12	47	1.3	925	CF825990	CF825990 EST703372
13	46.2	1.2	645	CNS01213	AL101589 Drosophila
14	45.8	1.2	881	CF821532	CF821532 EST698914
15	45.6	1.2	506	BP153627	BP153627 BP153627
16	45.6	1.2	652	CNS010969	CNS010969 WHE3878 D
17	45.4	1.2	753	CG855542	CG855542 ZMMB8C025
18	45.4	1.2	1101	CNS01720	AL107511 Drosophila
19	45.2	1.2	1087	BZ573831	BZ573831 mab2_3381
20	45.2	1.2	1101	CNS0175X	AL108460 Drosophila
21	44.4	1.2	584	AV922850	AV922850 AV922850
22	44.4	1.2	1101	CNS0175X	AL108460 Drosophila
23	44.4	1.2	2100	CL945925	CL945925 OaIFSB001
24	43.8	1.2	914	BF257966	BF257966 HVSMEF001

25	43.4	1.2	587	6	CA679773	CA679773 wim4.pk0
26	43.2	1.2	966	9	CNS02R3T	AL210098 Tetrarodon
27	43.2	1.2	971	9	CNS04VB9	AL308862 Tetrarodon
28	43.2	1.2	1019	7	CNS041IT	AL292286 Tetrarodon
29	43	1.1	487	7	CV039737	CV039737 4137143 B
30	43	1.1	540	6	CD727100	CD727100 4031393 1
31	43	1.1	573	6	CD216293	CD216293 PGP2n.pk0
32	43	1.1	580	1	AV922834	AV922834 AV922834
33	43	1.1	591	4	B1391789	B1391789 PGP1n.pk0
34	43	1.1	601	4	BM425602	BM425602 PGP2c.pk0
35	43	1.1	615	4	BM425623	BM425623 PGP2c.pk0
36	43	1.1	664	1	AJ450428	AJ450428 AJ450428
37	43	1.1	665	1	AJ455557	AJ455557 AJ455557
38	43	1.1	667	1	AJ450427	AJ450427 AJ450427
39	43	1.1	673	7	CO507218	CO507218 GGEZB201
40	43	1.1	674	1	AJ449106	AJ449106 AJ449106
41	43	1.1	690	1	AJ445947	AJ445947 AJ445947
42	43	1.1	695	4	BM486338	BM486338 pgm1c.pk0
43	43	1.1	698	5	BU481178	BU481178 603471787
44	43	1.1	704	6	CD763744	CD763744 GGEZLB100
45	43	1.1	727	6	CD764115	CD764115 GGEZLB100

#### ALIGNMENTS

RESULT 1  
CNS0091P  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION  
AL053013  
VERSION  
GI:4934461  
KEYWORDS  
GSS.  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Drosophila melanogaster  
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mamooser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

#### COMMENT

FEATURES  
source  
1. 925  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR19D16"  
/clone\_lib="RPCI-98"  
/note="end : TET3"

#### ORIGIN

Query Match 1.3%; Score 49.2; DB 9; Length 925;  
Best Local Similarity 13.9%; Pred. No. 0.34;

	Matches	51, Conservative	161, Mismatches	154, Indels	0, Gaps	0,
QY	761	TCTCTGGGGGAGATGGGTAGCCGCCCCCACTCGCTCCCCCAAGTCGTGACTTTCGATG	820			
Db	553	TTTSGSGTGGKCCSSGSGSBSCCSCSSCSGSCBCCCCSCSGSSSSSSSSSKSS	612			
QY	821	GCGCGCGCATTTGCTGTGTCGCGCTGTGGACATAGCGCTTGGGAAGGTGCTTGTGACA	880			
Db	613	TSBCTSCCCSSKSXVCTGTCSSSSSCSSSSSSSTSSSTSSSTSSKSSSGSSSSSSSYTS	672			
QY	881	TCTTGGCGGCTATGTAGACAGAGAGTGGCGCGCTGTGACCTTTAAGTCAATGACG	940			
Db	673	KSTASGSGSWASAGGSGSGTSTSSSSSSSTSSSSSVSGSKSTBSGSSBSGSSS	732			
QY	941	GCGAAATCCCTCCACCGAGACCTGTGTTACTTACTCCCGGCATCTCTCTCTGTG	1000			
Db	733	SSSTSSBSCTSTSSSSSSSYSSSTCSCTCCCSYSSSTSSSSSTSSWGTSGSSSV	792			
QY	1001	CCCTGATGTGTGGGGTCTGTGGCAGCAGATAGCTGTGCGCACATGTGATGAGAGAG	1060			
Db	793	GTSSSSDSTSTCCSGCCCTCTCCSTTBYMBCTTSSCGGSSSSSGAGVTKTGCGCGCSSS	852			
QY	1061	GAGCTGTGACGTGATGAACCGGCTGATAGCGTTGCTCGCGGGGTAAACATGTTTCC	1120			
Db	853	TNGMBGTSSACSSSSSSSCSSSSSVSSKSSASSSSVSSGSSGVSNSSSKSSSGSV	912			
QY	1121	CCAGCGC	1126			
Db	913	SSGSGS	918			

```

RESULT 2
CF824031/c
LOCUS          CF824031          552 bp      mRNA      linear      EST 01-APR-2004
DEFINITION    E87701413 Coccidioides posadasii saprobic phase cDNA library, 2 to
               4 kb Coccidioides posadasii cDNA clone CIDAP05 5' end, mRNA
               sequence.
ACCESSION     CF824031
VERSION       CF824031.1  GI:45930088
KEYWORDS
SOURCE
ORGANISM      Coccidioides posadasii
               Coccidioides posadasii
               Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
               Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE     1 (bases 1 to 552)
               Gardner,M.J. and Cole,G.T.
               Analysis of gene expression in Coccidioides posadasii mycelia and
               spherules via expressed sequence tags
               Unpublished (2003)
JOURNAL
COMMENT       Other ESTs: E8701412
               Contact: Gardner MJ
               The Institute for Genomic Research
               9712 Medical Center Drive, Rockville, MD 20850, USA
               Tel: 301 838 3519
               Fax: 301 838 0208
               Email: gardner@tigr.org
               Seq primer: M13 Reverse.
FEATURES
             location/Qualifiers
             1..552

```

```

/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDAP05"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"
/note="Vector: pEXpress 1; Site 1: Not I; Site 2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

```

	Query Match	Best Local Similarity	1.3%;	Score 48;	DB 7;	Length 552;
	Matches 117;	Conservative	0;	Mismatches 115;	Indels 0;	Gaps 0;
QY	797	CCAGTGTCTTAAAGCTTTCGTAGGCGCCGCGCATGTGTGGTGGCGCTTTGGCAGCATAG	856			
Db	522	CCATGGCCGCGATCTTGATGACCGCCGTGTGGCCATCGTCTTGCCCTCTGT	463			
QY	857	GCCTTGGAGAGTCTTGATGACATCTTGGCGGGCTATGAGCAGAGTGGACGCCGC	916			
Db	462	CCATGGCCGCGGTCTTGATGACATATGTCATAGCCGCGGTCTTGCGCGTGTGC	403			
QY	917	TCTGTGCTTTAAGTATATAGCGGCGAAATGCCCTTACACCGAGAACTGGTTAACTTAC	976			
Db	402	TTGTGGCCGCGTCTCATGACCGCCGTGTGTGGCCATCATGTTCATGCGCGCCGTGC	343			
QY	977	TCCCTGGCATCTCTCTCTGTGTGCCCTGTGTCGCGGGTCTGATGCGCAGC	1028			
Db	342	TTGTGGCCATCATGTTCATGACCGCCGCTCTTGTTGGCGCGTGTTCATATGCG	291			

RESULT 3	
CF824624/c	
LOCUS	
DEFINITION	856 bp mRNA linear EST 01-Apr-2004
ACCESSION	BF824624
VERSION	EST702006 Coccidioides posadasii saprobic phase cDNA library, 2 to
KEYWORDS	4 Kb Coccidioides posadasii cDNA clone CIDMS39 5' end, mRNA sequence.
SOURCE	CF824624
ORGANISM	CF824624.1 GI:45930681
REFERENCE	EST.
AUTHORS	Coccidioides posadasii
TITLE	Coccidioides posadasii
COMMENT	Eukaryote; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Omygenales; mitosporic Omygenales; Coccidioides.
JOURNAL	1 (bases 1 to 856)
COMMENT	Gardner,M.J. and Cole,G.T.
COMMENT	Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags
COMMENT	Unpublished (2003)
COMMENT	Contact: Gardner MJ
COMMENT	The Institute for Genomic Research
COMMENT	9712 Medical Center Drive, Rockville, MD 20850, USA
COMMENT	Tel: 301 838 3519
COMMENT	Fax: 301 838 0208
COMMENT	Email: gardner@tigr.org
FEATURES	Seq primer: M3 Reverse.
SOURCE	Location/Qualifiers
	1..856

Qy	797	CCAGTGTGCTTACAGCTTTGCTTGAAGCGCCGCGCATTTGCTGTGTGCGCGCTTTGGACAGCATAG	856
Db	791	CCATGCGCGCGCTACTTGTGTGACCCCGCGTGTGTGTGCGCATCGTACTTGTGCGCGCTCGGT	732
Qy	857	GCCTTGGGAAGTGTCTTGTGTGACATCTTTGGCGGCGCTATGTAGAGAGAGATGTGCAAGCGCGC	916
Db	731	CCATGTGCGCGCGTGTGTGTGTGTGCGCATCATGTGCATGTGCGCGCGCGTGTGTGTGCGCGCTGTGC	672

Query Match 1.3%; Score 48; DB 7; Length 856;  
 Best Local Similarity 50.4%; Pred. No. 0.6%;  
 Matches 117, Conservative 0; Mismatches 115; Indels 0; Gaps 0;

ORIGIN

```

/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/_clone="CIDAS39"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phase resistant"
/_clone_1ib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"
/_note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

```

QY 917 TGTGGCTTTAAGTCATGAGGCGGAATGCCCTCCACGAGACCTGTAACTAC 976  
 |||||  
 DB 671 TTGTGACCTCTGTCTCATGCGCCGCTGTTGTGACCATCATGTCATGCGCCGCTGC 612  
 |||||  
 QY 977 TCCCTGCATCTCTCTCTGTGTGTCCTGTGTCGCGGTGTGTGCGCAGC 1028  
 |||||  
 DB 611 TTGTGACCATCATGTCATGCGCGCGCTGTGTGCGCGCTGTGTGCGCATGCGC 560  
 |||||

RESULT 4  
 CNS0091P/c 925 bp DNA linear GSS 03-JUN-1999  
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #  
 DEFINITION BAC191016 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION  
 AL053013 GI:4934461  
 VERSION  
 AL053013.1  
 KEYWORDS  
 Drosophila melanogaster (fruit fly)  
 SOURCE  
 Drosophila melanogaster  
 ORGANISM  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
 Genoscope.  
 Direct Submision  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 Determination of this BAC-end sequence was carried out as part of a  
 collaboration with the Berkeley Drosophila Genome Project (BDGP).  
 The BDGP is constructing a physical map of the Drosophila  
 melanogaster genome using these BACs. For further information  
 please see <http://www.fruitfly.org> The BDGP Drosophila  
 melanogaster BAC library was prepared by Kazuo Osoegawa and  
 Aaron Mammeter in Pieter de Jong's laboratory in the Department of  
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
 NY. The library is named RPCL-98 and was constructed by partial  
 EcoRI digestion of Drosophila DNA provided by the BDGP from the  
 isogenic strain Y2; cn bw sp, the same strain used for the BDGP's  
 P1 and EST libraries. A more detailed description of the library  
 and how to order individual BAC clones, the entire library, or  
 filters for hybridization from the BACPAC Resource Center can be  
 found at [http://bacpac.med.buffalo.edu/drosophila\\_pac.htm](http://bacpac.med.buffalo.edu/drosophila_pac.htm).

FEATURES  
 source  
 1..925  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone\_lib="RPCL-98"  
 /note="end : TET3"

ORIGIN  
 Query Match 1.3%; Score 47.6; DB 9; Length 925;  
 Best Local Similarity 15.3%; Pred. No. 0.9; Mismatches 148; Indels 3; Gaps 1;  
 Matches 57; Conservative 164

QY 1798 GAGTCGACAGCTCCCATGCGAGCCGGAACGATGATGATGCTCACTTCATGCTAC 1857  
 |||||  
 DB 914 SSBSCGSSSMSTSS 855  
 |||||  
 QY 1858 GACCCCTCCACATCAGCAGAGCGGTAAGCGAGCTGCGCAGGGGCTCTCCCCC 1917  
 |||||  
 DB 854 NASSSGCGCGCAGABCCMCSSSSSCGASARGVKVRASGAGKRGCGGAGABASHSS 795  
 |||||  
 QY 1918 TCTTGGCGAGCTCTTACCTAGCCAGTTG---TCTGCCCTCTTCGAAAGGGAATAC 1974  
 |||||  
 DB 794 ACBSSSSSCSACWSASSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 735  
 |||||  
 QY 1975 ATTACCAAAATGACTTCCAGACGCTGACCTCATGAGGCAACTCCTGTGGCGGAT 2034  
 |||||

DB 734 SSSSCSSSVSCSVASMSCSBSSSSASASSSSSSSSSSSSSSSCSCCTSMSCSCTSA 675  
 |||||  
 QY 2035 GAGATGGGGGGAGACTTACCGCGTGAAGTCAAGAACAGATATATCTCGACTCT 2094  
 |||||  
 DB 674 SMSAARSS 615  
 |||||  
 QY 2095 TTGACACCCCTCCAGCGAGAGATGATGAGCCGGAATGTCGCTCCCGCGAGATCTG 2154  
 |||||  
 DB 614 SASGMSSSSVSSGCMSSSGCGGCGVGSSSSSSSSSSSSSSSSSSSSSSSSS 555  
 |||||

QY 2155 CGGAATCCAG 2166  
 |||||  
 DB 554 AAAAASCVAS 543  
 |||||

RESULT 5  
 CF824954 978 bp mRNA linear EST 01-APR-2004  
 LOCUS CF824954  
 DEFINITION EST702316 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb Coccidioides posadasii cDNA clone CIDA34.3' end, mRNA sequence.

ACCESSION  
 CF824954  
 VERSION  
 CF824954.1 GI:45931011  
 KEYWORDS  
 EST.  
 SOURCE  
 Coccidioides posadasii  
 ORGANISM  
 Coccidioides posadasii  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; Malesporici Onygenales; Coccidioides.

REFERENCE  
 1 (bases 1 to 978)  
 Gardner M.J. and Cole G.T.  
 Analysis of gene expression in Coccidioides posadasii mycelia and  
 spherules via expressed sequence tags  
 Unpublished (2003)  
 Other ESTs: EST702337  
 Contact: Gardner MJ  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301 838 3519  
 Fax: 301 838 0208  
 Email: gardner@cigr.org.

JOURNAL  
 Coccidioides posadasii  
 TITLE  
 Coccidioides posadasii  
 AUTHORS  
 Gardner M.J. and Cole G.T.  
 ANALYSIS  
 Analysis of gene expression in Coccidioides posadasii mycelia and  
 spherules via expressed sequence tags  
 Unpublished (2003)  
 Other ESTs: EST702337  
 Contact: Gardner MJ  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301 838 3519  
 Fax: 301 838 0208  
 Email: gardner@cigr.org.

FEATURES  
 source  
 1..978  
 Location/Qualifiers  
 /organism="Coccidioides posadasii"  
 /mol\_type="mRNA"  
 /strain="C735"  
 /db\_xref="taxon:199306"  
 /clone\_lib="CIDA34"  
 /dev\_stage="saprobic phase (mycelia)"  
 /lab\_host="E. coli DH10B, T1 phage resistant"  
 /clone\_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"  
 /note="Vector: pExpress 1; Site\_1: Not I; Site\_2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

ORIGIN  
 Query Match 1.3%; Score 47.6; DB 7; Length 978;  
 Best Local Similarity 50.0%; Pred. No. 0.91; Mismatches 119; Indels 0; Gaps 0;  
 Matches 119; Conservative 0

QY 783 CCAACTCGCTCCCGCAGTGTGCTTACAGCTTGTGAGGCGCGGATGCTGTGCGG 842  
 |||||  
 DB 732 CTTTCATCTTCATATGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 791  
 |||||  
 QY 843 TGTGGCAGCATAGGCTTGTGGAAGTCTTGTGACATCTTGTGCGGCGTATAGAGCAG 902  
 |||||  
 DB 792 TGTGGCGGT 851  
 |||||  
 QY 903 AGTGGAGGCGGCTGTGTGCTTTAAGGTATAGAGCGGGAAGGCTCCACGAGAGA 962  
 |||||  
 DB 852 TGTGGCGGT 911  
 |||||  
 QY 963 CCGTGTAACTTACTCCCTGATCTCTCTGTGTGCTGTGTGTGTGTGTGTGTGTGT 1020  
 |||||

[illegible]

REFERENCE	1
AUTHORS	Laurentino B.C., Ruiz J.C. and Cruz A.K.
TITLE	CGS analysis of the Leishmania braziliensis genome
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 635)
AUTHORS	Cruz A.K.
TITLE	Direct Submission
JOURNAL	Submitted (17-JUN-2003) Cruz A.K., University of Sao Paulo, Department of Molecular and Cell Biology, FMRP, Avenida Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL
COMMENT	Clone requests: akcruz@imp.usp.br.
FEATURES	Location/Qualifiers
SOURCE	1..635
ORIGIN	/organism="Leishmania braziliensis"
	/mol_type="genomic DNA"
	/strain="WHOM/BR/75/M2904"
	/db_xref="taxon:5660"
	/clone="LBAF018C06"
Query Match	1.3%; Score 47; DB 9; Length 635;
Best Local Similarity	48.7%; Pred. No. 1.2;
Matches 128; Conservative	0; Mismatches 135; Indels 0; Gaps 0;
QY	780 CGCCCACTGCGTCCCCAGATGCTGCTTGCATTTGATAGCGCCGCGCATTTGCTGTGC 839
DB	382 CGCCCACTTCAAGTTCCTGCTGTGACGCGAGAGTCCACTGCACCGCTGTGC 323
QY	840 GCGTGTGGCAGCATAGGCGCTTGGAGAGTCTTGTGACATCTTGGCGGGCTATGAGC 899
DB	322 TGCTCAGAGATCCATGACTTTCTCTCGTATGATCTGTATGACTCACCCTCAGCGGGA 263
QY	900 AGAGATGGCAGGCGCGCTCGTGGCTTTAAGTCATAGCGCGGAATGCCCTTCAACGA 959
DB	262 AGGATGCTGCTGCTGCTGCTCAAGCTTCTGTGAGACTCCGGAAACAGATCCCGACGA 203
QY	960 GGACCTGTAACTTATCTCCCTGCGATCTCTCTCTGAGCCCTGTGCTGCGGGGTCGT 1019
DB	202 GCGCGGTACATGCCCTGAGAACCTTCTGCTCCCATATGAGAGCTTGTCTGTCT 143
QY	1020 GTGCGCAGCAGTACTGCGTGGC 1042
DB	142 GTTGGCGCGGTGTCGAGACTGC 120
RESULT 8	
COO15154/c	855 bp mRNA linear EST 09-JUN-2004
LOCUS	EST78536 Coccidioides posadasii saprobic phase cDNA library,
DEFINITION	greater than 4kb Coccidioides posadasii cDNA clone C1B875 5' end,
ACCESSION	COO15154
VERSION	COO15154.1 GI:48522043
KEYWORDS	EST.
SOURCE	Coccidioides posadasii
ORGANISM	Coccidioides posadasii
	Eukaryote; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
	Oxyxales; mitosporic Oxyxales; Coccidioides.
REFERENCE	1 (bases 1 to 855)
AUTHORS	Gardner M.J. and Cole G.T.
TITLE	Analysis of gene expression in Coccidioides posadasii mycelia and
	spherules via expressed sequence tags
JOURNAL	Unpublished (2003)
COMMENT	Other ESTs: EST78535
	Contact: Gardner MJ
	The Institute for Genomic Research
	9712 Medical Center Drive, Rockville, MD 20850, USA
	Tel: 301 838 3519
	Fax: 301 838 0208
	Email: gardner@igr.org
	Seq primer: M13 Reverse.
FEATURES	Location/Qualifiers
SOURCE	1..855





Db 897 GTGCTGTGGCGCTGCTTCATGAGCCGCCGCTTGCCATCATGTCATGAGCCGCC 838  
| | | | |  
QY 898 GCAGAGTGCAGAGCCGCTCGTGCCTTTAGATGATGAGCGAGAAATGCCCTCCACC 957  
| | | | |  
Db 837 GTGCTTGTGGCCGCTCGTGTGCGCGCTGCTGCATGAGCCGCCGCTTGAGCCATC 778  
| | | | |  
QY 958 GAGAGCTGTGTTACTTACTCCCTGCCATCTCTTCCTGAGGCGCTGTCGAGGATC 1017  
| | | | |  
Db 777 ATGTTCATGAGCCCGCTTGTTGGCCATGTCATGATGAGCCGCCGCTTGAGCCGTC 718  
| | | | |  
QY 1018 GTGTGGGAGC 1028  
| | | | |  
Db 717 GTGTCCATGAGC 707  
| | | | |

RESULT 11  
CF823378/c 923 bp mRNA linear EST 01-APR-2004  
LOCUS EST700760 Coccidioides posadasii saprobic phase cDNA library, 2 to  
DEFINITION 4 kb Coccidioides posadasii cDNA clone CIDAL17 5' end, mRNA  
sequence.  
ACCESSION CF823378  
VERSION CF823378  
KEYWORDS EST.  
SOURCE GI:45929435  
ORGANISM Coccidioides posadasii  
Coccidioides posadasii  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Oryzales; mitosporic Oryzales; Coccidioides.  
REFERENCE 1 (bases 1 to 923)  
Gardner, M.J. and Cole, G.T.  
Analysis of gene expression in Coccidioides posadasii mycelia and  
spherules via expressed sequence tags  
Unpublished (2003)  
Other ESTs: EST700759  
JOURNAL Contact: Gardner MJ  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301 838 3519  
Fax: 301 838 0208  
Email: gardner@igr.org  
COMMENT Seq primer: M13 Reverse.  
LOCATION/Qualifiers  
1. .923  
/organism="Coccidioides posadasii"  
/mol\_type="mRNA"  
/strain="C735"  
/db\_xref="taxon:199306"  
/clone="CIDAL17"  
/dev\_stage="saprobic phase (mycelia)"  
/lab\_host="E. coli DH10B, T1 phage resistant"  
/clone\_lib="Coccidioides posadasii saprobic phase cDNA  
library, 2 to 4 kb"  
/note="Vector: pExpress 1, Site\_1: Not I; Site\_2: Eco RV;  
Coccidioides posadasii saprobic phase cDNA library, size  
fractionated cDNA 2 to 4 kb"

ORIGIN  
Query Match 1.3%; Score 47; DB 7; Length 923;  
Best Local Similarity 52.9%; Pred. No. 1.3;  
Matches 101; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 838 GCGGCTGTGGAGCATAGGCTTGGAGAGTGGTGTGACATCTTGGCGGCTATGGA 897  
| | | | |  
Db 907 GTGCTTGTGGCCGCTCGTGTGCGCGCTGTTGACATCATGTCATGAGCCGCC 848  
| | | | |  
QY 898 GCAGAGTGCAGAGCCGCTCGTGTGCGCTTTAAGTCAATGAGCGGAGAAATGCCCTCCACC 957  
| | | | |  
Db 847 GTGCTTGTGGCCGCTCGTGTGCGCGCTGTCATGAGCCGCCGCTTGAGCCATC 788  
| | | | |  
QY 958 GAGAGCTGTGTTACTTACTCCCTGCCATCTCTCTCTGTGAGCCCTGTCGAGGATC 1017  
| | | | |  
Db 787 ATGTTCATGAGCCCGCTTGTTGGCCATGTCATGATGAGCCGCCGCTTGAGCCGTC 728  
| | | | |

QY 1018 GTGTGGGAGC 1028  
| | | | |  
Db 727 GTGTCCATGAGC 717  
| | | | |

RESULT 12  
CF825990/c 925 bp mRNA linear EST 01-APR-2004  
LOCUS EST703372 Coccidioides posadasii saprobic phase cDNA library, 2 to  
DEFINITION 4 kb Coccidioides posadasii cDNA clone CIDB026 5' end, mRNA  
sequence.  
ACCESSION CF825990  
VERSION CF825990.1 GI:45932047  
KEYWORDS EST.  
SOURCE Coccidioides posadasii  
Coccidioides posadasii  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Oryzales; mitosporic Oryzales; Coccidioides.  
REFERENCE 1 (bases 1 to 925)  
Gardner, M.J. and Cole, G.T.  
Analysis of gene expression in Coccidioides posadasii mycelia and  
spherules via expressed sequence tags  
Unpublished (2003)  
Other ESTs: EST703371  
JOURNAL Contact: Gardner MJ  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301 838 3519  
Fax: 301 838 0208  
Email: gardner@igr.org  
COMMENT Seq primer: M13 Reverse.  
LOCATION/Qualifiers  
1. .925  
/organism="Coccidioides posadasii"  
/mol\_type="mRNA"  
/strain="C735"  
/db\_xref="taxon:199306"  
/clone="CIDB026"  
/dev\_stage="saprobic phase (mycelia)"  
/lab\_host="E. coli DH10B, T1 phage resistant"  
/clone\_lib="Coccidioides posadasii saprobic phase cDNA  
library, 2 to 4 kb"  
/note="Vector: pExpress 1, Site\_1: Not I; Site\_2: Eco RV;  
Coccidioides posadasii saprobic phase cDNA library, size  
fractionated cDNA 2 to 4 kb"

ORIGIN  
Query Match 1.3%; Score 47; DB 7; Length 925;  
Best Local Similarity 52.9%; Pred. No. 1.3;  
Matches 101; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 838 GCGGCTGTGGAGCATAGGCTTGGAGAGTGGTGTGACATCTTGGCGGCTATGGA 897  
| | | | |  
Db 919 GTGCTTGTGGCCGCTCGTGTGCGCGCTGTTGACATCATGTCATGAGCCGCC 860  
| | | | |  
QY 898 GCAGAGTGCAGAGCCGCTCGTGTGCGCTTTAAGTCAATGAGCGGAGAAATGCCCTCCACC 957  
| | | | |  
Db 859 GTGCTTGTGGCCGCTCGTGTGCGCGCTGTCATGAGCCGCCGCTTGAGCCATC 800  
| | | | |  
QY 958 GAGAGCTGTGTTACTTACTCCCTGCCATCTCTCTCTGTGAGCCCTGTCGAGGATC 1017  
| | | | |  
Db 799 ATGTTCATGAGCCCGCTTGTTGGCCATGTCATGATGAGCCGCCGCTTGAGCCGTC 740  
| | | | |  
QY 1018 GTGTGGGAGC 1028  
| | | | |  
Db 739 GTGTCCATGAGC 729  
| | | | |

RESULT 13  
CNS01213 645 bp DNA linear GSS 26-JUL-1999  
LOCUS CNS01213  
DEFINITION Drosophila melanogaster genome survey sequence 17 end of BAC  
BACN08C07 of DrosBAC library from Drosophila melanogaster (fruit

fly), genomic survey sequence.  
 ACCESSION AL101589  
 VERSION AL101589.1 GI:5613200  
 KEYWORDS GSS.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 645)  
 REFERENCE Genoscope.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobac11.  
 FEATURES  
 source Location/Qualifiers  
 1..645  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone\_1ib="BACN08C07"  
 /clone\_1ib="DrosBAC"  
 /plasmid="pBelobac11"  
 /note="end : T7"  
 ORIGIN  
 Query Match 1.2%; Score 46.2; DB 9; Length 645;  
 Best Local Similarity 12.0%; Pred. No. 1.9;  
 Matches 55; Conservative 127; Mismatches 276; Indels 1; Gaps 1;  
 QY 800 GTGCTCTTCAAGTCTTCTGAGGCGCGGCACTTGTGCGGCTGTGGACATAGGCC 859  
 DB 129 KKGAGGAGGNNNGTGTGTTGNNNGGGGNNNNNNNNNNNNNNNNNNNNNNNN 188  
 QY 860 TTGGAGAGTGTGTGACATCTTGGAGGCGGCTATGAGACAGAGTGGCGGCGTCG 919  
 DB 189 NNSCCSNCTCANNNTNTNTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 248  
 QY 920 TGGCCTTTAAGTCATGAGGCGGGAATGCCCTCCACCGAGACTGTGTTACTTCTCC 979  
 DB 249 NSGSS 308  
 QY 980 CTGCCATCTCTCTCTGTCGCTGCTGCGGGTCTGTGCGGACGATCTGCGTC 1039  
 DB 309 TSTTNTSNTSS 368  
 QY 1040 GGCACGTGGTCCAGGAGGAGGAGTGTGATGAACCGGCTGATAGCGCTGCGCT 1099  
 DB 369 NNSNATTTST -NGSSGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 427  
 QY 1100 CGGAGGATACCATGTTTCCCGACGACATATGTGCGAGAGCGCGCGACGACGTG 1159  
 DB 428 SSSSSSSSSSSSTTTTNSASNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 487  
 QY 1160 TCACTCAGATCTCTCCGACCTTACTATACCCCACTGTTGAAGAGGCTCCACAGTGA 1219  
 DB 488 SCWTTTANNTSSTNTTNTAGNNNATSSSSSSSSSSSSSSSSSSSSSSSS 547  
 QY 1220 TTAACGAGAGCTGCTCAGCGCCTGCTCGGCTCGTGGC 1258  
 DB 548 TGTBSSGNSNSNNNSCNSNNCTGNSNSSSSTTTGSS 586  
 RESULT 14  
 CF821532/c

LOCUS CF821532 881 bp mRNA linear EST 01-APR-2004  
 DEFINITION EST698914 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb Coccidioides posadasii cDNA clone CIDA65 5' end, mRNA sequence.  
 ACCESSION CF821532  
 VERSION CF821532.1 GI:45927589  
 KEYWORDS EST.  
 SOURCE Coccidioides posadasii  
 ORGANISM Coccidioides posadasii  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; Mitosporic Onygenales; Coccidioides.  
 1 (bases 1 to 881)  
 REFERENCE Gardner M.J. and Cole G.T.  
 AUTHORS Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags  
 JOURNAL Unpublished (2003)  
 COMMENT Other ESTs: EST698913  
 CONTACT: Gardner MJ  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301 838 3519  
 Fax: 301 838 0208  
 Email: gardner@tigr.org  
 Seq primer: M13 Reverse.  
 FEATURES  
 source Location/Qualifiers  
 1..881  
 /organism="Coccidioides posadasii"  
 /mol\_type="mRNA"  
 /strain="C735"  
 /db\_xref="taxon:199306"  
 /clone="CIDA65"  
 /dev\_stage="saprobic phase (mycelia)"  
 /lab\_host="E. coli DH10B, TI phage resistant"  
 /clone\_1ib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"  
 /note="Vector: pEXpress 1; Site 1: Not 1; Site 2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"  
 ORIGIN  
 Query Match 1.2%; Score 45.8; DB 7; Length 881;  
 Best Local Similarity 55.3%; Pred. No. 2.6;  
 Matches 89; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
 QY 868 GTGCTGTGACATCTTGGCGGCGCTATGAGACAGAGTGGCGGCGCTGTGGCTTT 927  
 DB 876 GTGCTGTGACATCTTGGCGGCGCTATGAGACAGAGTGGCGGCGCTGTGGCTTT 917  
 QY 928 AAGTCATGAGGCGGGAATGCCCTCCACGAGACTGTGTTACTTCTGCTGCTC 987  
 DB 816 GTGTCATGAGCGCGCTGTGCTGTCGATGATGATGATGATGATGATGATGATG 757  
 QY 988 CTCTCTCTGTCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1028  
 DB 756 ATGTCATGAGCGCGCGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 716  
 RESULT 15  
 BP153627  
 LOCUS BP153627/c 506 bp mRNA linear EST 30-DEC-2003  
 DEFINITION BP153627 full-length enriched swine cDNA library, adult ovary sus scrofa cDNA clone OVM10094G12 5', mRNA sequence.  
 ACCESSION BP153627  
 VERSION BP153627.1 GI:40403100  
 KEYWORDS EST.  
 SOURCE Sus scrofa (pig)  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 1 (bases 1 to 506)  
 REFERENCE Uenishi H., Eguchi T., Suzuki K., Sawazaki T., Toki D., Shinkai H., Okumura N., Hamasima N. and Awata T.  
 PEDF (pig EST Data Explorer): construction of a database for Ests

JOURNAL  
COMMENT

derived from porcine full-length cDNA libraries  
Nucleic Acids Res. 32 (1), D484-D488 (2004)  
Contact: Hirohide Uenishi  
Animal Genome Laboratory, Genome Research Department  
National Institute of Agrobiological Sciences  
2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan  
Tel: +81-29-838-8627  
Fax: +81-29-838-8627  
Email: huenishi@affrc.go.jp

EST project with full-length enriched cDNA libraries carried out in  
Animal Genome Research Program (Japan) by National Institute of  
Agrobiological Sciences and STAFF-Institute  
Single pass sequencing of clones derived from oligo-capped cDNA  
library

Vector sequences were eliminated by RepeatMasker version 2002/07/13  
and crossmatch version 0.990319  
Low quality bases were trimmed based on the quality values.

## FEATURES

## source

Location/Qualifiers  
1..506  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/clone="ORNL0094G12"  
/tissue\_type="ovary"  
/dev\_stage="adult"  
/clone\_lib="full-length enriched swine cDNA library, adult  
ovary"

## ORIGIN

```

Query Match      1.2%; Score 45.6; DB 5; Length 506;
Best Local Similarity 48.5%; Pred. No. 2.6;
Matches 126; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 2234 AGGCCCCGGACTACGTCCTCCCTGAGTGTAGATGGTCCCACTGCCACTTAAGACC 2293
    |||||
DB 389 AGCCGCTGGGCTGCTCTCTCCAGCGCTCTTGACCTTAAGTGTATGAATCAACGCTGCTT 330
    |||||

QY 2294 CTCTATACCACTCCAGGAGAGAGACAGTTGTTCTGACAGAAATCCACCGTGTCTT 2353
    |||||
DB 329 TGCTGCTGGGCGGCAACCGCTCCGGGAGCGAGGCCAGGCCCGGCCCTTTCCCG 270
    |||||

QY 2354 CTGCCCTGGGGAGCTTGCCAAAGGCTTTGGTAGCTCCGAACCGTGGCCGTCGACA 2413
    |||||
DB 269 CAGCGCAGGCGGAGCTGCGGAGCGACGAGACCGAGCCGAGCCGAGCCGAGCGGCGCT 210
    |||||

QY 2414 GCGGCAAGGCAACGCCCTCTCTGACCACTTCGACGACGCGGAGCAGATCTGACG 2473
    |||||
DB 209 CGGCTCGCGCAGGAGGCGCTGCAAGGCCCAACCCGAAAGCGCGGCAACCGGTCGCG 150
    |||||

QY 2474 TTGATCGTATTCCTCCATG 2493
    |||||
DB 149 GGGCGCGAGGGCCACCAAG 130
    |||||

```

Search completed: February 21, 2005, 18:48:32  
Job time : 11369 secs